

The δN formula is the dynamical renormalization group

Article (Submitted Version)

Dias, Mafalda, Ribeiro, Raquel H and Seery, David (2013) The δN formula is the dynamical renormalization group. *Journal of Cosmology and Astroparticle Physics*, 2013 (10). pp. 1-45. ISSN 1475-7516

This version is available from Sussex Research Online: <http://sro.sussex.ac.uk/id/eprint/48791/>

This document is made available in accordance with publisher policies and may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher's version. Please see the URL above for details on accessing the published version.

Copyright and reuse:

Sussex Research Online is a digital repository of the research output of the University.

Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable, the material made available in SRO has been checked for eligibility before being made available.

Copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

The δN formula is the dynamical renormalization group

Mafalda Dias,¹ Raquel H. Ribeiro,^{2,3} and David Seery¹

¹ Astronomy Centre, University of Sussex, Falmer, Brighton, BN1 9QH, UK

² Department of Physics, Case Western Reserve University, 10900 Euclid Ave, Cleveland, OH 44106, USA

³ Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Wilberforce Road, Cambridge CB3 0WA, UK

E-mail: M.Dias@sussex.ac.uk, RaquelHRibeiro@case.edu, D.Seery@sussex.ac.uk

Abstract. We derive the ‘separate universe’ method for the inflationary bispectrum, beginning directly from a field-theory calculation. We work to tree-level in quantum effects but to all orders in the slow-roll expansion, with masses accommodated perturbatively. Our method provides a systematic basis to account for novel sources of time-dependence in inflationary correlation functions, and has immediate applications. First, we use our result to obtain the correct matching prescription between the ‘quantum’ and ‘classical’ parts of the separate universe computation. Second, we elaborate on the application of this method in situations where its validity is not clear. As a by-product of our calculation we give the leading slow-roll corrections to the three-point function of field fluctuations on spatially flat hypersurfaces in a canonical, multiple-field model.

Keywords: inflation, cosmology of the very early universe, cosmological perturbation theory, non-gaussianity

Contents

1 Introduction

Since Maldacena’s computation of the three-point function produced by an epoch of single-field inflation [?], considerable effort has been invested in obtaining correlation functions for more complex scenarios. In part this investment has been motivated by the increasing sophistication of observational cosmology. But an equally important influence has been the drive to understand the implications of our inflationary theories, especially in examples with many interacting degrees of freedom.

The tool used to extract predictions from inflationary scenarios is quantum field theory, in which each observable is constructed from correlation functions. Quantum field theory incorporates the effect of fluctuations on all scales, making each correlation function potentially sensitive to both short- and long-distance effects. Over very short distances this implies that they must be defined carefully by a renormalization prescription, as in flat spacetime. But the most interesting outcome of our investment in studying correlation functions has been the realization that they have a rich infrared structure which is very different to the case of flat spacetime. Some aspects of this structure were reviewed in Refs. [? ? ?].

Nontrivial structure in correlation functions is often associated with the existence of kinematic hierarchies. A paradigmatic example from particle physics is a hadronic jet carrying some large energy E but small invariant mass m_{jet} . Another is electroweak processes at large momentum transfer $s \gg M_Z^2$. (Here s is the Mandelstam variable and M_Z is the mass of the Z boson.) The kinematic region where ratios such as E/m_{jet} or s/M_Z^2 become large is called the *Sudakov region*. In this region logarithms of large ratios invalidate perturbation theory, and to obtain even a qualitative description requires resummation.

Analogues of these effects are responsible for the rich infrared structure of inflationary correlation functions. Viewed as a sum of Feynman diagrams, a generic n -point function is characterized by the comoving momenta carried by each propagator. At tree-level these are linear combinations \mathbf{q}_i of the external momenta $\{\mathbf{k}_1, \dots, \mathbf{k}_n\}$. Another scale is the comoving Hubble length $\tau \sim (aH)^{-1}$ where a is the scale factor and $H = \dot{a}/a$ is the Hubble rate. Sudakov-like logarithms involving ratios of these scales appear in each correlation function.¹

When there are no hierarchies among the $|\mathbf{q}_i|$ and τ , each of these scales will be comparable to some reference scale k_* . If the n -point functions are expressed in terms of background quantities evaluated at the horizon-crossing time for k_* , the Sudakov-like logarithms are small and the correlation functions are simple. However, if large hierarchies exist it will not be possible to find a k_* for which all logarithms are small and the correlation functions develop a complex structure. Indeed, we will see that it is necessary to carry out a resummation, just as for hadronic jets and high-energy electroweak observables. Understanding the details of this enhanced structure is particularly important because correlation functions with large hierarchies are valuable observables: for example, n -point functions with $q_i/q_j \gg 1$ can be

¹Despite the appealing analogy, the logarithms which appear in inflationary correlation functions are *not* Sudakov logarithms in the following sense: in scattering calculations the Kinoshita–Lee–Nauenberg theorem implies that sufficiently inclusive observables are infrared safe, due to cancellation between virtual diagrams and real soft or collinear emission [? ?]. Sudakov effects arise in comparatively exclusive observables because exclusion of phase space regions prevents complete cancellation. The Sudakov logarithms measure this large remainder. In inflation the background fields themselves have time dependence and there is no meaningful sense in which these logarithms arise as a cancellation between real and virtual effects. For this reason we shall refer to them as *Sudakov-like logarithms*.

used to determine nonlinear and stochastic halo bias [?]. But despite their significance, a systematic treatment of these effects has not appeared in the literature.

In this paper we begin a systematic treatment of correlation functions containing Sudakov-like logarithms. We focus on logarithms containing the comoving Hubble scale τ , which may all be rewritten in the form $\ln|k_*\tau|$ (perhaps at the cost of introducing other logarithms of the form $\ln q_i/k_*$). In this form they coincide with the ‘time-dependent’ logarithms $\ln|k_*\tau|$ described in Ref. [?]; see also Ref. [?]. We will return to the remaining ‘scale’ and ‘shape’ logarithms (involving $\ln q_i/k_*$ and $\ln q_i/q_j$, respectively [?]) in a future publication. We work to tree-level in quantum effects but to *all orders* in the slow-roll expansion. In particular, this means that masses are accommodated perturbatively. We will comment on the implications of this choice as we develop our argument in §§??–?? below.

Separate-universe formulae. It has been understood for some time that logarithms of the form $\ln|k_*\tau|$ describe time evolution of each n -point function [? ? ? ?]. But, although the necessity to account for time evolution is well-known, the precise role of the $\ln|k_*\tau|$ terms has received comparatively little attention. Most discussion of time dependence are framed in terms of the ‘separate universe’ approach to perturbation theory [? ? ? ?]. On very large scales this gives a description of a perturbed universe by patching together unperturbed solutions with different initial conditions. One implementation follows Sasaki & Tanaka [?], who argued that the superhorizon limit of perturbation theory could be obtained from Jacobi fields of the background phase space. (A similar interpretation was recently given in Ref.[?].) Related discussions of perturbation theory on large scales were given in Refs. [? ? ?]. On the basis of these arguments one can determine the evolution of correlation functions by tracking the evolution of these Jacobi fields and averaging over a suitable ensemble of stochastic initial conditions.

This approach gives a reliable description of effects occurring on superhorizon scales, but it cannot describe subhorizon physics or phenomena which occur near the epoch of horizon exit. For example, this leads to some ambiguity about the precise ensemble of initial conditions which should be used when computing correlation functions, to which we will return in §??. In this paper we argue that the separate-universe method can be reproduced by resummation of Sudakov-like logarithms. Among other benefits, this provides precise information about the role of horizon-crossing effects which cannot be captured from the superhorizon limit of perturbation theory.

After constructing suitable Jacobi fields and setting up an ensemble of initial conditions $\delta\phi_i(\mathbf{k}_1)$, the separate-universe approach we have just described asserts that the two- and three-point functions for a set of light scalar fields during inflation satisfy [?]

$$\langle\delta\phi_\alpha(\mathbf{k}_1)\delta\phi_\beta(\mathbf{k}_2)\rangle \supseteq \frac{\partial\phi_\alpha}{\partial\phi_i^*}\frac{\partial\phi_\beta}{\partial\phi_j^*}\langle\delta\phi_i(\mathbf{k}_1)\delta\phi_j(\mathbf{k}_2)\rangle_* \quad (1.1a)$$

$$\begin{aligned} \langle\delta\phi_\alpha(\mathbf{k}_1)\delta\phi_\beta(\mathbf{k}_2)\delta\phi_\gamma(\mathbf{k}_3)\rangle &\supseteq \frac{\partial\phi_\alpha}{\partial\phi_i^*}\frac{\partial\phi_\beta}{\partial\phi_j^*}\frac{\partial\phi_\gamma}{\partial\phi_k^*}\langle\delta\phi_i(\mathbf{k}_1)\delta\phi_j(\mathbf{k}_2)\delta\phi_k(\mathbf{k}_3)\rangle_* \\ &+ \frac{\partial^2\phi_\alpha}{\partial\phi_i^*\partial\phi_j^*}\frac{\partial\phi_\beta}{\partial\phi_k^*}\frac{\partial\phi_\gamma}{\partial\phi_m^*}\int\frac{d^3q}{(2\pi)^3}\langle\delta\phi_i(\mathbf{k}_1-\mathbf{q})\delta\phi_k(\mathbf{k}_2)\rangle_*\langle\delta\phi_j(\mathbf{q})\delta\phi_m(\mathbf{k}_3)\rangle_* \\ &+ \text{cyclic permutations.} \end{aligned} \quad (1.1b)$$

We have adopted the notation of Refs. [? ?], which will be used throughout this paper. On the left-hand side, each correlation function is evaluated at some late time t and transforms as a tensor in the tangent space associated with the field-space coordinate $\phi_\alpha(t)$. Indices in this tangent space are labelled $\{\alpha, \beta, \dots\}$. On the right-hand side each correlation function is evaluated at the horizon-crossing time t_* for the reference scale k_* , and transforms as a tensor in the tangent space associated with the field-space coordinate $\phi_i(t_*)$. Indices in this tangent space are labelled $\{i, j, \dots\}$. Where the field-space manifold is curved it is essential to preserve this distinction [?]. In this paper we will work with a flat field-space metric except in §?? but this index convention remains convenient. The bitensors $\partial\phi_\alpha/\partial\phi_i^*$, $\partial^2\phi_\alpha/\partial\phi_i^*\partial\phi_j^*$ measure variation of the inflationary trajectory under a change of initial conditions at the horizon-crossing time for the reference scale, and the symbol ‘ \supseteq ’ indicates that subleading corrections from ‘loop’ diagrams have been ignored together with disconnected contributions.

Eqs. (??)–(??) make a number of strong assertions. First, Eq. (??) asserts that, at all times and accounting for large contributions from all orders in the slow-roll expansion, the two-point function can be written as a linear combination of its values at some arbitrary earlier time with coefficients $\partial\phi_\alpha/\partial\phi_i^*$. This implies that (at least when decaying modes have become negligible) all time-dependent Sudakov-like logarithms can be factorized into these coefficients. Here and below, ‘to all orders in the slow-roll expansion’ is a statement about a perturbative expansion in terms of slow-roll parameters evaluated at the horizon-crossing time for k_* .

Second, Eq. (??) asserts that any time-dependent Sudakov-like logarithms appearing in the three-point function exhibit a similar factorizable structure. In principle the three-point function could have arbitrary dependence on the external momenta k_i [? ?]. Therefore, at each order, the Sudakov-like logarithms could enter with different functions of the k_i . But Eq. (??) asserts that, to all orders in the slow-roll expansion, almost all Sudakov-like logarithms factorize into the coefficients $\partial\phi_\alpha/\partial\phi_i^*$ which appear in the two-point function (??). Exceptions are allowed only for Sudakov-like logarithms which multiply functions of the k_i obtainable from a product of two-point functions. For massless fields, or where masses are taken into account perturbatively, these are the ‘local’ momentum combinations $(k_2^3 k_3^3)^{-1}$ and its permutations. Even for these exceptions, Eq. (??) requires that the resummation of Sudakov-like logarithms produces a result which is related in a nontrivial way to the factorizable coefficients already present in the two-point function.

Each of these assertions is a complex statement, applicable to all orders in the slow-roll expansion, about the structure of infrared divergences which can be produced by the underlying quantum field theory. Although these results are implied by the perturbation-theory results described above, it is nontrivial to see how they are reproduced by the Sudakov-like logarithms which arise from the underlying quantum field theory. In this paper we explain how this occurs by deducing the time-dependent structure of each correlation function directly from the Sudakov-like divergences it contains. We will see that this provides a practical and convenient means to compute the evolution of each n -point function in cases where the separate universe picture is complicated to apply, or an untrustworthy guide for our intuition.

Summary. The Sudakov-like time-dependent logarithms in inflationary correlation functions are qualitatively similar to those occurring in particle physics, and therefore can be understood using the same tools. In §?? we sketch the main steps, focusing on a physical

interpretation of the ‘factorizable’ structure. In §?? we begin to develop this argument in detail. The strategy breaks into two parts.

- First, in §?? we prove a *factorization theorem* which provides constraints on the terms in each correlation function which can be logarithmically enhanced. This factorization theorem plays the same role as a proof of renormalizability in applying renormalization-group arguments to the ultraviolet behaviour of correlation functions: it guarantees that all large logarithms can be assembled into a finite number of ‘renormalized’ functions.
- Second, in §??, we derive a *renormalization group equation* (‘RGE’) which can be used to determine these unknown ‘renormalized’ functions. Once the renormalization group equations have been solved the full correlation function can be reconstructed. It will turn out that the RGEs are equivalent to the transport equations used to determine $\partial\phi_\alpha/\partial\phi_i^*$ and their higher derivatives [? ? ? ? ? ?].

In §?? we illustrate our method using examples. When justified from the superhorizon limit of perturbation theory the separate-universe approach requires a set of initial conditions, usually assumed to be generated from quantum fluctuations which become classical after passing outside the horizon. Implicit in this point of view is a ‘matching’ between quantum and classical calculations. Several authors have noted that the precise choice of matching time is ambiguous, which could lead to uncertainties in an accurate calculation [? ? ?]. In §?? we revisit this question using the RGE approach, in which there is no need to invoke a matching from quantum to classical evolution because the entire calculation takes place within the framework of quantum field theory.² We show how this gives a unique prescription which resolves the matching ambiguity.

In §?? we illustrate the practical utility of the RGE method by using it to determine evolution equations for covariant versions of the coefficients $\partial\phi_\alpha/\partial\phi_i^*$ and their higher-order generalizations in the presence of a nontrivial field-space metric. These coefficients are not straightforward to determine using classical intuition, making the renormalization-group approach simple and reliable. This example could be adapted to more complex cases, perhaps including the effect of loop corrections, where the usual classical separate-universe arguments do not function or are prohibitively difficult to apply.

In §?? we conclude with a short discussion of our main results. We provide details of the next-order calculation of the three-point correlation function (which is required to deduce an appropriate renormalization group equation for the coefficient functions) in Appendix ??.

Notation and conventions. Throughout this paper, we adopt units in which $c = \hbar = 1$. Our index conventions for scalar fields are described below Eqs. (??)–(??). We work with the action

$$S = \frac{1}{2} \int d^4x \sqrt{-g} \left\{ M_{\text{P}}^2 R - \partial_a \phi_\alpha \partial^a \phi_\alpha - 2V \right\}, \quad (1.2)$$

where $V = V(\phi)$ is an arbitrary potential and Latin indices a, b, \dots , are Lorentz indices contracted with the spacetime metric g_{ab} . Except where explicitly indicated, scalar field indices are contracted using the flat field-space metric $\delta_{\alpha\beta}$. We will generally set the reduced Planck mass $M_{\text{P}} \equiv (8\pi G)^{-1/2}$ to unity.

²Note that this is purely a statement about the calculation of correlation functions. The need to understand how fluctuations decohere (presumably solving the Schrödinger’s cat problem) remains.

2 Why is resummation necessary?

It is not at all clear from inspection of Eqs. (??)–(??) that they involve a resummation of time-dependent Sudakov-like logarithms. In this section and the next we explain why resummation is necessary and develop a heuristic approach to these equations. In §?? we will consider the same problem from a more formal viewpoint, that of the dynamical renormalization group.

To quadratic order, the action for fluctuations of the scalar fields ϕ_α in (??), measured on uniform-expansion hypersurfaces, can be written

$$S = \frac{1}{2} \int d^3x dt a^3 \left(\delta\dot{\phi}_\alpha \delta\dot{\phi}_\alpha - \frac{1}{a^2} \partial_i \delta\phi_\alpha \partial_i \delta\phi_\alpha - m_{\alpha\beta} \delta\phi_\alpha \delta\phi_\beta \right), \quad (2.1)$$

where an overdot denotes differentiation with respect to cosmic time. The mass-mixing matrix $m_{\alpha\beta}$ can be written explicitly in terms of slow-roll parameters [? ? ?]

$$m_{\alpha\beta} = V_{\alpha\beta} - \frac{1}{a^3} \frac{d}{dt} \left(\frac{a^3}{H} \dot{\phi}_\alpha \dot{\phi}_\beta \right) = V_{\alpha\beta} - 3\dot{\phi}_\alpha \dot{\phi}_\beta - \varepsilon \dot{\phi}_\alpha \dot{\phi}_\beta - \frac{2}{H} \dot{\phi}_\alpha \ddot{\phi}_\beta, \quad (2.2)$$

where $\varepsilon = -\dot{H}/H^2$ is the usual slow-roll parameter. Inflation occurs whenever $\varepsilon < 1$. At lowest order in slow-roll the mass-matrix can be related to the ‘expansion tensor’ $u_{\alpha\beta}$ introduced in Ref. [?],

$$u_{\alpha\beta} = -\frac{m_{\alpha\beta}}{3H^2}. \quad (2.3)$$

Two-point function. It is convenient to work in conformal time, defined by $\tau = \int_\infty^t dt'/a(t')$. For the two-point function only a single hierarchy can be generated, measured by the ratio of k to the comoving Hubble scale $aH \approx -\tau^{-1}$. Therefore $|k\tau| \approx 1$ at horizon crossing. Working to next-order in the slow-roll expansion and evaluating the two-point function at least a little after horizon-crossing we find

$$\begin{aligned} \langle \delta\phi_\alpha(\mathbf{k}_1) \delta\phi_\beta(\mathbf{k}_2) \rangle_\tau &= (2\pi)^3 \delta(\mathbf{k}_1 + \mathbf{k}_2) \frac{H_*^2}{2k^3} \\ &\times \left\{ \delta_{\alpha\beta} \left[1 + 2\varepsilon_* \left(1 - \gamma_E - \ln \frac{2k}{k_*} \right) \right] + 2u_{\alpha\beta}^* \left[2 - \gamma_E - \ln(-k_*\tau) - \ln \frac{2k}{k_*} \right] + \dots \right\}, \end{aligned} \quad (2.4)$$

where ‘ \dots ’ indicates higher-order corrections which we have not computed. We have set $k = k_1 = k_2$ to be the common magnitude of the external momenta and dropped contributions which decay like positive powers of τ . The scale k_* determines a horizon-crossing time around which we have chosen to Taylor-expand background time-dependent quantities such as H . Evaluation at the horizon-crossing time for k_* is denoted by a subscript ‘ $*$ ’. At this stage its precise assignment is arbitrary and can be chosen to suit our own convenience. The terms involving $\ln k/k_*$ and $\ln(-k_*\tau)$ are our first examples of Sudakov-like logarithms.

This result has appeared in various forms in the literature. It was given in the single-field case by Stewart & Lyth [?], who worked directly in terms of the conserved comoving-gauge curvature perturbation \mathcal{R} . The corresponding result for field fluctuations in the uniform-expansion gauge was given by Nakamura & Stewart [?], who accounted for multiple fields and a nontrivial field-space metric. Higher-order corrections were given by Gong & Stewart,

who developed an algorithmic approach to compute them using Green’s functions of the Mukhanov–Sasaki equation [? ?].³

Callan–Symanzik equation. The reference scale k_* is not physical, and therefore the correlation functions cannot depend on it. Invariance under a change in k_* is expressed by the Callan–Symanzik equation. Taking all background quantities to be functions of the fields ϕ_α , this equation can be written

$$\left(\frac{\partial}{\partial \ln k_*} + \frac{\partial \phi_\lambda^*}{\partial \ln k_*} \frac{\partial}{\partial \phi_\lambda^*} \right) \langle \delta \phi_{\alpha_1}(\mathbf{k}_1) \cdots \delta \phi_{\alpha_n}(\mathbf{k}_n) \rangle = 0. \quad (2.5)$$

The differential operator in brackets (\cdots) is simply the total derivative $d/d \ln k_*$. Specializing (??) to the two-point function yields, to lowest-order in a slow-roll expansion,

$$\frac{d \ln H_*}{d \ln k_*} + \varepsilon_* = 0. \quad (2.6)$$

As above, a subscript ‘*’ indicates evaluation at the horizon-crossing point for k_* . Since (??) holds for any k_* it expresses the condition $\varepsilon = -\dot{H}/H^2$ which is satisfied for any change of the background fields ϕ_α which themselves satisfy the background equations of motion. This is sometimes expressed by saying that the change should be ‘on-shell’. Eq. (??) guarantees that we are free to change k_* provided that we compensate by adjusting all background quantities according to their classical equations of motion. The appearance of classical equations is a consequence of our restriction to tree-level processes.⁴

Time evolution. We now return to Eq. (??) and consider its behaviour for different values of k and τ .

An expression such as (??) is said to be computed using *fixed order* perturbation theory, because the calculation is carried to a predetermined order in the slow-roll expansion. When k is comparable to aH , Eq. (??) shows that we can choose k_* to be approximately their common magnitude provided we evaluate background quantities near the horizon-crossing time for k . Then $\ln(-k_*\tau)$ and $\ln k/k_*$ will both be negligible, and a fixed-order expression such as Eq. (??) is a good approximation. It is dominated by its lowest-order term $H_*^2/2k^3$.

More than a few e-folds after horizon-crossing the hierarchy $k/aH = |k\tau|$ becomes exponentially small. No matter how we choose k_* it is no longer possible to make both $\ln k/k_*$ and $\ln(-k_*\tau)$ negligible. Therefore we must accept the appearance of large, growing contributions which cannot be absorbed into background quantities by a choice of evaluation time. Growing terms of this kind are potentially hazardous. They are sometimes described as divergences, because on its own $\ln(-k_*\tau)$ becomes unboundedly large as $\tau \rightarrow 0$. This is

³Eq. (??) of this paper agrees with Eq. (43) of Gong & Stewart [?], although the time-dependent terms appear different due to the way Gong & Stewart selected their time of evaluation.

Results restricted to a two-field model were quoted by Byrnes & Wands [?] and Lalak et al. [?], who worked with an explicit adiabatic–isocurvature basis. Their results neglect the time-dependent term $\ln(-k_*\tau)$ and therefore disagree with our Eq. (??). Formulae compatible with (??) were given by Avgoustidis et al. [?].

⁴One might have expected to recover the equation of motion $3H\dot{\phi}_\alpha = -V_{,\alpha}$ for each field. This does not happen because (??) does not depend on the fields individually, but only through their aggregate contribution to H . Eq. (??) can be regarded as the equation of motion for H . In a model with a single field ϕ it is equivalent to the classical equation of motion $d\phi/dN = -\sqrt{2\varepsilon}$.

simply an artefact of perturbation theory, in the same way that the Taylor expansion of any function

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \dots$$

appears to diverge when $|x - x_0| \gg 1$. In reality, when $\varepsilon_* \ln(-k_*\tau) \sim 1$ we must sum an infinite number of terms from a fixed-order expression such as (??) before we can determine even its qualitative behaviour. Therefore, unless some principle or symmetry enforces a precise cancellation, fixed-order perturbation theory can never provide an adequate description. This obligation to deal with an infinite sequence of terms of the form $[\varepsilon_* \ln(-k_*\tau)]^n$ was emphasized by Weinberg [?], who went on to speculate that the methods of the renormalization group could be used to perform the resummation, in analogy with the case of Sudakov effects in QCD. In this paper we show that this is indeed the case.

We conclude that any expression which is valid to arbitrarily late times, such as (??), must resum an infinite number of terms. For example, consider the well-studied model of double quadratic inflation [? ? ? ?]. The action is

$$S = -\frac{1}{2} \int d^4x \sqrt{-g} \left\{ R + \partial_a \varphi \partial^a \varphi + \partial_a \chi \partial^a \chi + m_\varphi^2 \varphi^2 + m_\chi^2 \chi^2 \right\}, \quad (2.7)$$

where φ and χ are two light scalar fields with $m_\varphi, m_\chi < H$. We focus on the dimensionless two-point function $\Sigma_{\alpha\beta}$, defined by

$$\langle \delta\phi_\alpha(\mathbf{k}_1) \delta\phi_\beta(\mathbf{k}_2) \rangle_\tau = (2\pi)^3 \delta(\mathbf{k}_1 + \mathbf{k}_2) \frac{\Sigma_{\alpha\beta}}{2k^3}, \quad (2.8)$$

but similar remarks apply to higher n -point functions.

In Fig. ?? we present a typical super-horizon evolution of $\Sigma_{\varphi\varphi}$ calculated using the formula (??). There is a clearly defined peak in the vicinity of $N = -\ln(-k_*\tau) \simeq 20$, caused by a turn of the trajectory in field space [?], after which $\Sigma_{\varphi\varphi}$ asymptotes to a constant. Inflation is ongoing through the entire evolution, with ε growing to a maximum value ~ 0.26 near the peak. Fig. ?? shows clearly that $\Sigma_{\alpha\beta}$ can exhibit very rapid evolution. The linear approximation involving $\varepsilon_* \ln(-k_*\tau)$ is already poor for $N \gtrsim 10$ and is totally incorrect (even qualitatively) for $N \gtrsim 20$. If the initial time had been chosen closer to the peak, the linear approximation could easily have been invalidated within $\lesssim 1$ e-fold.

The flat asymptote for $N \gtrsim 20$ is associated with convergence to an ‘adiabatic limit’ in which the uniform-density gauge curvature perturbation ζ is conserved [?]. To describe this asymptotic plateau certainly requires resummation of an infinite number of powers of $N = -\ln(-k_*\tau)$. In any finite combination the highest power of N must eventually dominate, leading to power-law growth at large N . It is only in an infinite sum that sufficiently many terms remain available at high-order to balance the growing contribution at low orders, enabling $\Sigma_{\varphi\varphi}$ to be constant at late times.

3 Towards a resummation prescription

We have identified growth of the logarithmic ‘divergence’ in (??) with the infrared region where $k/aH \ll 1$. It represents the cumulative effect of interactions which operate over very many Hubble times. We would like to separate the physics of these interactions from the creation of fluctuations in each k -mode, which takes place over only a few Hubble times.

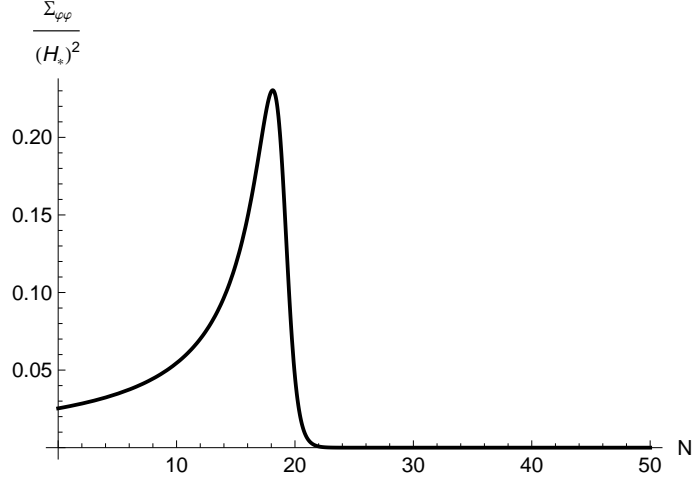


Figure 1: Super-horizon evolution of the two-point function $\Sigma_{\varphi\varphi}$ in the double-quadratic model (??), normalized by the Hubble scale at horizon crossing. This evolution is for the scale which left the horizon (at $N = 0$) fifty e-folds before the end of inflation (at $N = 50$). The initial conditions were $\varphi_{\text{init}} = 8.2M_{\text{P}}$ and $\chi_{\text{init}} = 12.9M_{\text{P}}$ with $m_{\varphi}/m_{\chi} = 9$.

The disparity of these timescales implies that, when they are created, the fluctuations do not depend on details of the inflationary model in which they are embedded. It is the systematic separation of short-distance, model-independent ‘hard’ physics and long-distance, model-dependent ‘soft’ effects which produces the factorization in Eqs. (??)–(??).

In this section we illustrate how to perform and interpret factorization of the two-point function, and indicate how to extend the method to higher n -point functions.

3.1 Factorization of the two-point function

We work with the dimensionless two-point function $\Sigma_{\alpha\beta}$ defined in Eq. (??). The formula (??) includes terms at lowest-order in the slow-roll expansion and $\mathcal{O}(\varepsilon)$ corrections. We describe these, respectively, as ‘lowest-order’ (LO) and ‘next-lowest order’ (or next-order, NLO) terms. Working to NLO, we identify

$$\Sigma_{\alpha\beta}^{\text{NLO}} \supseteq H_*^2 \left(\delta_{\alpha\beta} + 2r_{\alpha\beta}^* - 2u_{\alpha\beta}^* \ln(-k_*\tau) - 2M_{\alpha\beta}^* \ln \frac{2k}{k_*} \right). \quad (3.1)$$

The constant matrices $r_{\alpha\beta}$ and $M_{\alpha\beta}$ are

$$r_{\alpha\beta}^* = \varepsilon_* \delta_{\alpha\beta} (1 - \gamma_{\text{E}}) + u_{\alpha\beta}^* (2 - \gamma_{\text{E}}) \quad (3.2a)$$

$$M_{\alpha\beta}^* = \varepsilon_* \delta_{\alpha\beta} + u_{\alpha\beta}^*. \quad (3.2b)$$

Heuristic treatment. In what follows, we assume that $\Sigma_{\alpha\beta}$ contains at most powers of $\ln k$, but not powers of k itself. This is a consequence of approximate scale invariance. Our aim is to separate the long- and short-distance contributions to (??). It was explained above that when $k/aH \ll 1$ we cannot make a choice of k_* which eliminates all large logarithms. In effect, we do not have enough free scales to absorb every large contribution: it is only possible to use k_* to eliminate contributions from *one* of $\ln k/k_*$ and $\ln(-k_*\tau)$. We choose $k_* = k$ to eliminate $\ln k/k_*$ and introduce a new, arbitrary ‘factorization scale’ k_{F} . (However, we

continue to denote quantities evaluated at the horizon-crossing time for k_* with a subscript ‘*’.) Ultimately, we will see that the factorization scale describes the boundary between long- and short-distance physics. We write

$$\begin{aligned}\Sigma_{\alpha\beta}^{\text{NLO}} &\supseteq H_*^2 \left(\delta_{\alpha\beta} + 2\tilde{r}_{\alpha\beta}^* - 2u_{\alpha\beta}^* \left[\ln \frac{k}{k_F} + \ln(-k_F\tau) \right] \right) \\ &= H_*^2 \left(\delta_{\alpha i} - u_{\alpha i}^* \ln(-k_F\tau) \right) \left(\delta_{\beta j} - u_{\beta j}^* \ln(-k_F\tau) \right) \left(\delta_{ij} + 2\tilde{r}_{ij}^* - 2u_{ij}^* \ln \frac{k}{k_F} \right),\end{aligned}\tag{3.3}$$

where $\tilde{r}_{\alpha\beta} = r_{\alpha\beta} - M_{\alpha\beta} \ln 2$. In the first line k_F cancels out. The second line is valid to next-order, which is the accuracy to which we are working. Based on the structure of (??), we introduce a ‘form-factor’ $\Gamma_{\alpha i}$, which is a function only of τ and the factorization scale k_F ,

$$\Sigma_{\alpha\beta} = H_*^2 \Gamma_{\alpha i} \Gamma_{\beta j} \left(\delta_{ij} + 2\tilde{r}_{ij}^* - 2u_{ij}^* \ln \frac{k}{k_F} \right).\tag{3.4}$$

We will see that this form-factor modifies the lowest-order functional dependence of $\Sigma_{\alpha\beta}$ on background quantities, giving the two-point function an enhanced structure which is not visible in fixed-order perturbation theory. Modifications of this type are an inevitable consequence of large contributions which cannot be absorbed into a change of evaluation point for the background quantities already present in a fixed-order expression such as (??) or (??).

Since $\Sigma_{\alpha\beta}$ must be independent of k_F it is possible to write a Callan–Symanzik equation similar to (??). For (??) to be independent of k_F requires

$$\frac{d\Gamma_{\alpha i}}{d \ln k_F} = -\Gamma_{\alpha j} u_{ji}.\tag{3.5}$$

In Eqs. (??)–(??) we have made use of the index convention described in §??. The indices α, β, \dots label the tangent space at the late time τ , whereas indices i, j, \dots label the tangent space at the horizon-crossing time for the factorization scale k_F . Note that in (??) we have rewritten the time of evaluation for u_{ij} as this horizon-crossing time. This is formally acceptable because the error we incur is $\mathcal{O}(\varepsilon^2)$ and therefore below the precision to which we are working. In §?? below we will see how to give a more mathematically satisfactory formulation of this argument. In this section the discussion is not intended to be rigorous, but to highlight the physical meaning of factorization.

Eq. (??) is the ‘backwards’ evolution equation introduced by Yokoyama et al. [? ? ?] for the separate-universe coefficient $\partial\phi_\alpha/\partial\phi_i$. This discussion shows that it can equally be regarded as a kind of renormalization-group equation for the ‘Sudakov-like’ form-factor $\Gamma_{\alpha i}$. Its solution requires an initial condition. This can be deduced from (??), which gives $\Gamma_{\alpha i} = \delta_{\alpha i}$ when $|k_F\tau| \sim 1$. In combination with (??) the initial condition enables us to give a physical interpretation of the factorization procedure (see Fig. ??). Our starting point is a fixed-order expression such as (??).

We begin with $|k_F\tau| \sim 1$. With this choice the form-factors $\Gamma_{\alpha i} = \delta_{\alpha i}$ are trivial and contain no information, whereas the fixed-order piece involving $\ln k/k_F$ accounts for contributions from both long- and short-distance effects. These are effects arising from the regions $k/aH \gtrsim 1$ and $k/aH \lesssim 1$, respectively. We now use (??) to systematically move k_F closer to k . This adjustment moves all effects generated up to the point where $k_F/aH \lesssim 1$ into $\Gamma_{\alpha i}$. Effects generated when $k_F/aH \gtrsim 1$ remain within the fixed-order expression. Since (??)

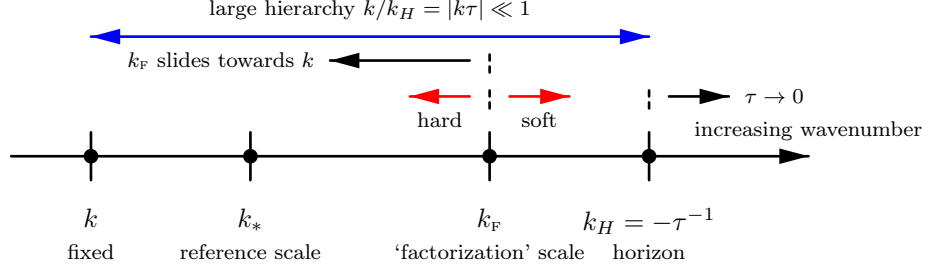


Figure 2: Scales appearing in factorization of the two-point function. At late times $\tau \rightarrow 0$, generating a large hierarchy between the horizon size $k_H = -\tau^{-1}$ and the fixed comoving scale k . The floating scale k_F acts as a boundary between the ‘hard’ (short-distance, $k/aH \gtrsim 1$) and ‘soft’ (long-distance, $k/aH \lesssim 1$) parts of the calculation. Beginning with $k_F = k_H$, all soft effects are included in the fixed-order part of the calculation. The ‘renormalization group equation’ (??) allows us to slide k_F towards k , successively absorbing long-distance effects in the Sudakov-like form-factor $\Gamma_{\alpha i}$. Eventually, only ‘hard’ effects dominated by comoving momenta $\sim k$ are included in the fixed-order part of the calculation.

merely expresses that $\Sigma_{\alpha\beta}$ is independent of k_F this does not change its numerical value. We terminate this process when $k_F \sim k$, at which point all soft effects generated when $k/aH \lesssim 1$ are absorbed into the Sudakov-like form-factor. Conversely, the hard factor in (??) receives contributions only from times when $k/aH \sim 1$. Because this hard factor is uncontaminated by any other scale,⁵ fixed-order perturbation theory can be used to obtain a good approximation.

Infrared safety. In the special case that the Sudakov-like logarithms vanish order-by-order there are no soft effects to absorb in $\Gamma_{\alpha i}$, and $\Gamma_{\alpha i} \approx \delta_{\alpha i}$ even for $k_F \sim k$. In these circumstances, the hard factor automatically receives contributions only from $k/aH \sim 1$. In Ref. [?] correlation functions with this property were described as ‘infrared safe’, by analogy with the same situation in scattering calculations. Infrared safe quantities decouple from the complex, cumulative interactions of soft modes and depend only on whatever physics dominates the hard subprocess. Observables built from such correlation functions have the advantage of being theoretically ‘clean’, but the disadvantage that they probe phenomena operating only over a narrow window of wavenumbers.

This discussion makes clear that the form-factor $\Gamma_{\alpha i}$ is the differential coefficient $\partial\phi_\alpha/\partial\phi_i$, and therefore that the factorization scale k_F corresponds to the initial time ‘*’ used to construct the separate universe formulae (??)–(??). Its arbitrariness corresponds to the arbitrary location of this initial time-slice. The hard, fixed-order contribution corresponds to the two-point function $\langle\delta\phi^\alpha\delta\phi^\beta\rangle_*$. As we explained above, this receives contributions only from times when $k \sim aH$ and describes the hard subprocess by which fluctuations are created. The creation process is rapid in comparison with the subsequent evolution, so separation of scales implies that its form is universal. All details of this discussion have precise parallels in the factorization of Sudakov-like form-factors in QCD [? ? ? ?].

3.2 Factorization of n -point functions with $n \geq 3$

With this background it is possible to consider higher n -point functions. We begin with the three-point function.

⁵For this statement to be strictly correct we must choose k_* to remove all other sources of large logarithms, in the same way that in scattering calculations we must choose the renormalization scale to do the same.

Removal of external wavefunctions. For reasons that will become clear below, it is helpful to discuss the n -point functions separately from their external wavefunction factors. In scattering calculations we sometimes discuss ‘amputated’ diagrams, meaning that the entirety of each external propagator is stripped away. In a nonequilibrium or time-dependent setting this cannot be done because half of the propagator participates in a nontrivial integral over the temporal position of the vertex to which it is attached. In Appendix ?? we show that the Feynman propagator takes the form⁶ [cf. Eq. (??)]

$$\langle T \delta\phi_\alpha(\mathbf{k}_1, \tau) \delta\phi_\beta(\mathbf{k}_2, \tau') \rangle = (2\pi)^3 \delta(\mathbf{k}_1 + \mathbf{k}_2) \times \begin{cases} w_{\alpha\gamma}^*(k, \tau) w_{\gamma\beta}(k, \tau') & \tau < \tau' \\ w_{\alpha\gamma}(k, \tau) w_{\gamma\beta}^*(k, \tau') & \tau' < \tau \end{cases}, \quad (3.6)$$

where $k = |\mathbf{k}_1| = |\mathbf{k}_2|$ is the common magnitude of the momenta on each external line, the operator T denotes time ordering and the matrix-valued mode functions $w_{\alpha\beta}$ corresponding to each half of the propagator are contracted with each other. In Fig. ?? we depict the conjugated mode $w_{\alpha\beta}^*$ as a dashed half-line, and the unconjugated mode $w_{\alpha\beta}$ by a solid half-line. The index contraction is denoted by a cross joining the dashed and solid halves. With these conventions, the three-point function can be depicted as in Fig. ?. The external, conjugated mode functions (dashed lines) are evaluated at some time τ which is taken to be later than the time η associated with the internal vertex. Finally in Fig. ?? we show the 3-point function with these external, conjugated wavefunction factors removed. The rules of the ‘in-in’ formulation of quantum field theory (see Appendix ??) show that the full three-point function should be obtained by adding Fig. ?? and its complex conjugate. Therefore the three-point function has the structure

$$\langle \delta\phi_\alpha(\mathbf{k}_1) \delta\phi_\beta(\mathbf{k}_2) \delta\phi_\gamma(\mathbf{k}_3) \rangle_\tau = 2 \operatorname{Re} \left[w_{\alpha\lambda}^*(k_1, \tau) w_{\beta\mu}^*(k_2, \tau) w_{\gamma\nu}^*(k_3, \tau) I_{\lambda\mu\nu}(\tau, k_1, k_2, k_3) \right], \quad (3.7)$$

where the vertex integral $I_{\lambda\mu\nu}$ corresponds to Fig. ?? and is determined using the methods of the in-in formalism. It represents the cumulative amplitude for three-body interactions up to time τ , and is given by an integral over the internal wavefunctions (which measure the probability of three particles interacting at a point), weighted by geometrical factors (which measure the volume in which the interaction can take place) and terms representing the detailed structure of each three-body interaction.

External wavefunction divergences. The discussion in §?? shows that although each wavefunction $w_{\alpha\beta}(k, \tau)$ contains ‘divergent’ logarithms as $\tau \rightarrow 0$, these are absorbed into the form-factor $\Gamma_{\alpha i}$ in the combination $\Gamma_{\alpha j}(\tau, k_F) w_{ji}(k, -k_F^{-1})$. (Here, the wavefunction w_{ji} is evaluated at the horizon-crossing time for the factorization scale k_F .) In concrete terms, although the ‘divergent’ logarithms in $w_{ij}(k, -k_F^{-1})$ now appear as powers of $\ln k_F$, the combination $\Gamma_{\alpha j}(\tau, k_F) w_{ji}(k, -k_F^{-1})$ is independent of k_F . Therefore these ‘divergent’ logarithms cancel. Meanwhile, the overall τ -dependence is determined by the solution to the renormalization group equation (??) and does not suffer from the appearance of ‘divergent’ terms. The same applies to $w_{\alpha\beta}^*$ provided we arrange for the phase of $w_{\alpha\beta}$ to be constant at late times.

⁶Computation of expectation values requires the ‘in-in’ or ‘Schwinger’ formalism, in which all field degrees of freedom are doubled and therefore there are four possible time-ordered two point functions. Strictly, Eq. (??) represents the time-ordered ++ two point function. For details, see Appendix ?? or Refs. [? ?].

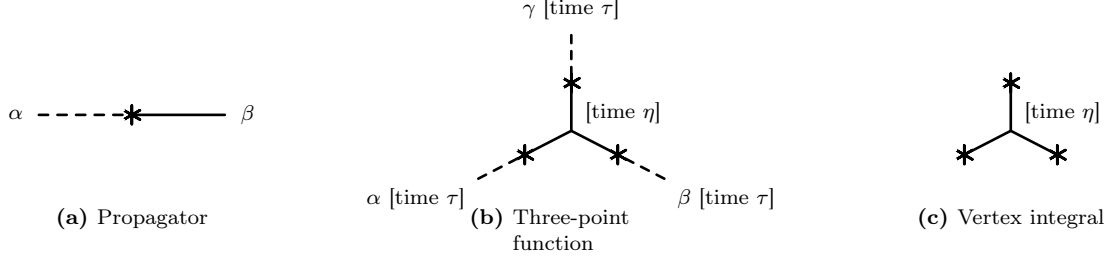


Figure 3: Structure of the three-point function as a ‘vertex integral’ over individual wavefunction factors. Fig. ?? depicts the propagator, which is a product of one conjugated wavefunction (dashed line) and one unconjugated wavefunction (solid line). The cross denotes contraction of these factors by matrix multiplication. In Fig. ??, three ‘internal’ factors participate in an integral over the vertex whereas the three ‘external’ factors do not. Note that the time at the vertex, η , is always earlier than the external time τ . In Fig. ?? we show the vertex integral with these external factors removed.

We now generalize this to the three-point function. Consider the combination

$$\Gamma_{\alpha i}(\tau, k_F) \Gamma_{\beta j}(\tau, k_F) \Gamma_{\gamma k}(\tau, k_F) \langle \delta\phi^i(\mathbf{k}_1) \delta\phi^j(\mathbf{k}_2) \delta\phi^k(\mathbf{k}_3) \rangle_{-k_F^{-1}} \quad (3.8)$$

which appears as one component of the separate-universe formula (??). Comparison with Fig. ?? and Eq. (??) shows that, by construction, we expect each form-factor Γ to absorb the ‘divergent’ $\ln k_F$ terms from its corresponding external wavefunction. But if the vertex integral $I_{\lambda\mu\nu}$ also contributes $\ln k_F$ terms these can not be absorbed by the Γ factors. Such large, ‘divergent’ contributions would remain, leaving residual k_F -dependence in (??).

Eq. (??) introduces a new form-factor $\partial^2\phi_\alpha/\partial\phi_i^*\partial\phi_j^*$, which provides a means by which these $\ln k_F$ terms can be absorbed. However, as explained in §??, it imposes stringent conditions (to all orders in the slow-roll expansion) on the structure of any $\ln k_F$ terms which appear in the vertex integral $I_{\lambda\mu\nu}$ —these must all be proportional to the momentum combination $k_1^{-3}k_2^{-3}$ (or its permutations) which can be generated by a product of two-point functions. Our key task in applying the renormalization-group formalism to the three-point function will be to prove a ‘factorization theorem’ which guarantees that $I_{\lambda\mu\nu}$ *does* possess this structure.

Higher n -point functions. Essentially the same discussion can be given for the four- and higher n -point functions, for which the separate universe method gives a structure similar to (??). Each n -point function contains a term like (??) [? ?], which absorbs $\ln k_F$ contributions from the external wavefunctions. Once this is done, a finite number of form-factors are available to absorb $\ln k_F$ contributions coming from integration over the vertices of the diagram, analogous to the vertex integral $I_{\lambda\mu\nu}$. The number of possible form-factors typically increases with n , although they are not all independent.

For example, for the four-point function we can write

$$\langle \delta\phi_\alpha(\mathbf{k}_1) \delta\phi_\beta(\mathbf{k}_2) \delta\phi_\gamma(\mathbf{k}_3) \delta\phi_\delta(\mathbf{k}_4) \rangle_\tau = 2 \operatorname{Re} \left[w_{\alpha\lambda}^*(k_1, \tau) w_{\beta\mu}^*(k_2, \tau) w_{\gamma\nu}^*(k_3, \tau) w_{\delta\pi}^*(k_4, \tau) I_{\lambda\mu\nu\pi}(\tau, \mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4) \right], \quad (3.9)$$

where the ‘vertex’ integral $I_{\lambda\mu\nu\pi}$ receives contributions at tree-level from graphs with the two topologies shown in Fig. ??. As above, factors of $\ln k_F$ from the external wavefunctions of both diagrams can be absorbed by contracting with $\Gamma_{\alpha i}$. Comparison with the formulae

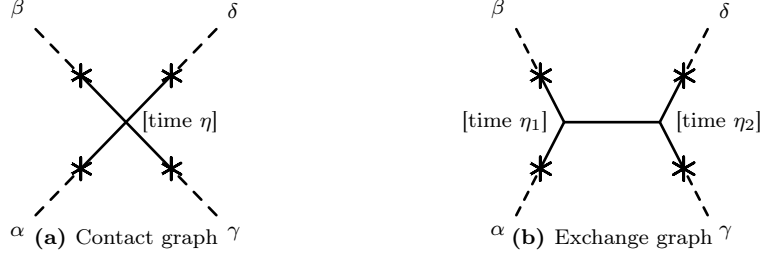


Figure 4: Form-factors available for the four-point function.

of Refs. [?] shows that form-factors are available to absorb factors of $\ln k_F$ from $I_{\lambda\mu\nu\pi}$ which are proportional to $k_1^{-3}k_2^{-3}k_3^{-3}$ or $k_1^{-3}k_2^{-3}|\mathbf{k}_1 + \mathbf{k}_3|^{-3}$ (or their permutations). These two possibilities correspond to the contact and exchange graphs, respectively. However, the form-factor for $k_1^{-3}k_2^{-3}|\mathbf{k}_1 + \mathbf{k}_3|^{-3}$ is built out of the form-factor used to absorb divergences in the three-point function. Only the form-factor for the contact interaction of Fig. ?? is ‘new’. A similar pattern repeats for all higher n , with only the irreducible n -point contact graph generating ‘new’ divergences: all others must be absorbed by form-factors which have already appeared in m -point functions with $m < n$. It is clear from inspection of Fig. ?? and its analogues for higher n -point functions that this arrangement is plausible because all diagrams except the contact graph are built out of lower-order vertices whose divergences can be described in terms of lower-order form-factors.

4 The dynamical renormalization group

In this section we revisit the analysis of §§??–?? from a slightly more formal viewpoint—that of the ‘dynamical renormalization group’. The original purpose of these methods was to determine dynamical scaling laws for correlation functions near an out-of-equilibrium critical point [? ?]. However, they have found various applications to inflationary correlation functions [? ? ? ? ?]. The approach developed here is similar to the discussion of time-dependence generated from loop corrections given by Burgess et al. [?]. More recently, Collins et al. suggested that the dynamical renormalization group method could be used to determine time evolution generated by integrating out heavy modes [?]. Our analysis demonstrates the steps which would be required to achieve this for an arbitrary n -point function.

Role of ‘renormalizability’. To apply the renormalization group requires a guarantee that all large logarithms can be absorbed into a finite number of ‘renormalized’ quantities. When applied to ultraviolet behaviour this guarantee is provided either by the criterion of renormalizability (in a strictly renormalizable theory), or the fact that only a finite number of irrelevant operators need to be kept in order to make predictions at fixed accuracy (in an effective field theory).

In the present case there is no notion of ‘renormalizability’ and we do not have either of these guarantees. The property which replaces renormalizability is factorization, in the sense of §§??–??. In particular, what is required is a guarantee that all divergences produced by ‘vertex integrals’ such as $I_{\lambda\mu\nu}$ and $I_{\lambda\mu\nu\pi}$ are proportional to a finite number of combinations of the external momenta. As we saw in §??, the separate universe asserts that all

divergences generated by $I_{\lambda\mu\nu}$ produce the shape $k_1^{-3}k_2^{-3}$ (or its permutations), and all divergences generated by $I_{\lambda\nu\nu\pi}$ produce the shapes $k_1^{-3}k_2^{-3}k_3^{-3}$ or $k_1^{-3}k_2^{-3}|\mathbf{k}_1 + \mathbf{k}_3|^{-3}$ (or their permutations).

These properties (and their analogues for higher n -point functions) are implied by the superhorizon structure of perturbation theory, but it is difficult to see them emerge at the level of Feynman diagrams. A rigorous proof is a ‘factorization theorem’. Theorems of this kind provide the missing guarantee that only a finite number of form-factors can be used to absorb all large logarithmically-enhanced contributions. They are an integral part of the apparatus used to study perturbative QCD processes such as jets and deep inelastic scattering [? ? ? ? ?].

4.1 Two-point function

In this section our purpose is to prove a factorization theorem for the vertex integral $I_{\lambda\mu\nu}$. Before doing so, we briefly return to Eqs. (??) and (??) and repeat the analysis of §?? in a way which can be generalized more easily to the three-point function.

Ignoring terms which decay like positive powers of $|k\tau|$, the coefficient $\Sigma_{\alpha\beta}$ defined in (??) depends on k and τ only through powers of $\ln(-k_*\tau)$ and $\ln k/k_*$. Therefore it can be interpreted as a Taylor series in these logarithms, expanded around the arbitrary horizon-crossing time for k_* . The renormalization group is a method to reverse-engineer a function from the first few terms in its Taylor expansion. When applied to $\Sigma_{\alpha\beta}$ this allows an all-orders reconstruction from information about the lowest terms in the perturbative series. The term ‘dynamical’ merely indicates that the ordinary renormalization-group method is being applied to a series expansion in time rather than energy.

Dynamical renormalization group analysis. To proceed we define a ‘hard’ contribution to $\Sigma_{\alpha\beta}$ which excludes any enhancement due to soft effects from $\ln|k_*\tau|$,

$$\Sigma_{\alpha\beta}^{\text{hard}*} = H_*^2 \left(\delta_{\alpha\beta} + 2r_{\alpha\beta}^* - 2M_{\alpha\beta}^* \ln \frac{2k}{k_*} \right). \quad (4.1)$$

We interpret $\Sigma_{\alpha\beta}^{\text{hard}*}$ as a function of k and k_* , but not τ . The full two-point function can be written

$$\Sigma_{\alpha\beta} \simeq \Sigma_{\alpha\beta}^{\text{hard}*} - \left[\Sigma_{\alpha\lambda}^{\text{hard}*} u_{\lambda\beta}^* + \Sigma_{\beta\lambda}^{\text{hard}*} u_{\lambda\alpha}^* \right] \ln(-k_*\tau), \quad (4.2)$$

where ‘ \simeq ’ indicates that this expression is valid up to NLO accuracy. In particular, we have replaced $H_*^2 \delta_{\alpha\beta}$ with $\Sigma_{\alpha\beta}$ in the factor multiplying $\ln(-k_*\tau)$. Although there is a potential mismatch in this exchange, it would appear at *next-next-order* and is therefore below the accuracy to which we are working. In the solution to the renormalization-group equation this will translate to an error below leading-logarithmic accuracy.⁷

⁷ The leading-logarithm approximation resums terms of the form $[\varepsilon_* \ln(-k_*\tau)]^n$ for all n , but not $\varepsilon[\varepsilon_* \ln(-k_*\tau)]^n$ or smaller. For a renormalization group equation of the schematic form

$$-\frac{\partial \Sigma}{\partial \ln \tau} = u \Sigma, \quad (4.3)$$

it can be proved by standard methods (see, eg., Ref. [?]) that resummation of the leading logarithms requires knowledge of u to $\mathcal{O}(\varepsilon)$; next-to-leading logarithms requires knowledge of u to $\mathcal{O}(\varepsilon^2)$, and so on. Mathematically, our methods will justify the separate universe method to leading-logarithm order. Although this is likely to be sufficient for observable inflation, subleading logarithms must become important over very long timescales.

We now interpret the right-hand side of (??) as a Taylor series expansion around the arbitrary horizon-crossing time for k_* . According to Taylor's theorem,

$$\Sigma_{\alpha\beta}(k, \tau) = \Sigma_{\alpha\beta}|_{|k_*\tau|=1} + \left. \frac{d\Sigma_{\alpha\beta}}{d\ln\tau} \right|_{|k_*\tau|=1} \ln(-k_*\tau) + \dots, \quad (4.4)$$

and so the coefficient of $\ln(-k_*\tau)$ is the derivative of $\Sigma_{\alpha\beta}$ evaluated at $|k_*\tau| = 1$. Therefore

$$-\left. \frac{d\Sigma_{\alpha\beta}}{d\ln\tau} \right|_{|k_*\tau|=1} = \Sigma_{\alpha\lambda}^{\text{hard}*} u_{\lambda\beta}^* + \Sigma_{\beta\lambda}^{\text{hard}*} u_{\lambda\alpha}^* = \Sigma_{\alpha\lambda}^* u_{\lambda\beta}^* + \Sigma_{\beta\lambda}^* u_{\lambda\alpha}^*. \quad (4.5)$$

A priori this is a statement about the value of the derivative only at the fixed time $\tau = -k_*^{-1}$. But precisely because k_* is arbitrary, the function $d\Sigma_{\alpha\beta}/d\ln\tau|_{|k_*\tau|=1}$ obtained in this way must be the same function of the horizon-crossing time for k_* as $d\Sigma_{\alpha\beta}/d\ln\tau$ is of τ . Therefore we can immediately promote it to a differential equation valid for all τ , obtained by removing ‘*’ from all terms in (??)

$$-\frac{d\Sigma_{\alpha\beta}}{d\ln\tau} = \Sigma_{\alpha\lambda} u_{\lambda\beta} + \Sigma_{\beta\lambda} u_{\lambda\alpha}. \quad (4.6)$$

In this equation $\Sigma_{\alpha\beta}$ is a function of k and τ , but $u_{\alpha\beta}$ is a function of τ only. In the literature it is often rewritten in terms of the inflationary e-folding time $N = \int H dt$ using $-d\ln\tau = dN$.

It might appear that (??) could have been obtained by differentiation of Eq. (??) with respect to τ . This procedure would produce an equation which is symbolically identical, but where the right-hand side is evaluated at $|k_*\tau| = 1$. To understand why (??) is valid for arbitrary τ it is necessary to construct (??) via comparison with Taylor's theorem, as described above. This justifies the more heuristic method used to obtain Eq. (??).

An explicit solution of Eq. (??) requires an initial condition. According to (??) we can obtain this initial condition from the constant term in the Taylor expansion, $\Sigma_{\alpha\beta}^{\text{hard}*}$.^{8,9} To obtain an accurate estimate using Eq. (??) we should set $k_* = k$ to remove large contributions from powers of $\ln k/k_*$, which yields¹⁰

$$\Sigma_{\alpha\beta} = H_k^2 \left(\delta_{\alpha\beta} + 2r_{\alpha\beta}^k - 2M_{\alpha\beta}^k \ln 2 + \dots \right) \quad \text{at } |k\tau| = 1, \quad (4.7)$$

where ‘...’ indicates higher-order corrections beyond NLO, which we have not computed, and a sub- or superscript ‘ k ’ now indicates evaluation at the horizon-crossing time for k . Together, Eqs. (??), (??) and (??) provide an interpretation of the renormalization-group analysis. The

⁸Therefore only the lowest two Taylor terms are required to reconstruct $\Sigma_{\alpha\beta}$. However, these receive contributions from all orders in the slow-roll expansion.

⁹In any renormalization-group procedure there is some arbitrariness in setting the initial condition (??), because we are free to treat some or all of the constant term $\Sigma_{\alpha\beta}^{\text{hard}*}$ as an additive constant rather than the zero-order term in the Taylor expansion. This makes no difference in an exact calculation, but can influence the result at finite orders in perturbation theory. A prescription for this choice is called a *factorization scheme*. For details, see eg. Ref. [?]. We absorb the entirety of $\Sigma_{\alpha\beta}^{\text{hard}*}$ as the initial condition.

¹⁰At $|k\tau| = 1$ there are contributions to $\Sigma_{\alpha\beta}$ from decaying power-law corrections which we have not written explicitly. (See also §??.) These terms make no contribution to (??) because they do not form part of the Taylor series at late times.

The contribution from $M_{\alpha\beta}$ could be removed by choosing $k_* = 2k$. For higher n -point functions, the analogous contribution would be removed by choosing $k_* = k_t$. To maximize the accuracy of each initial condition one should therefore evolve each n -point function from a marginally different initial time, but for the three- and four-point functions this effect will be negligible.

hard subprocess—here, creation of fluctuations near horizon exit—depends only on the scale k , and therefore fixed-order perturbation theory can be used to obtain an accurate estimate such as (??). The absence of other scales implies we can be confident that the subleading terms represented by ‘...’ are smaller than those we have calculated. This subprocess is used as an initial condition for the renormalization-group evolution, which evolves it to lower energies—here, represented by the decreasing value of k/aH . This evolution is accomplished by successively ‘dressing’ the hard factor with the cumulative effect of soft processes.

Eq. (??) is the dynamical renormalization group equation for the two-point function. It has already appeared in the literature as a ‘transport’ equation for $\Sigma_{\alpha\beta}$, as described by Mulryne et al. [? ? ?]. If desired, the methods of Ref. [?] can be used to convert (??) into an equivalent equation for the form-factor $\Gamma_{\alpha i}$ which coincides with (??). Therefore, the approach of this section is exactly equivalent to that of §??, or the standard δN formula [?].

4.2 A factorization theorem for the bispectrum

Let us return to the factorization properties of $I_{\lambda\mu\nu}$. We should interpret divergences in this vertex integral to mean that three-body interactions continue to arbitrarily late times, rather than being localized to the neighbourhood of horizon-crossing. Therefore the divergences originate from a kinematic region where all decaying modes become extinct. In this region the wavefunctions $w_{\alpha\beta}$ lose their \mathbf{k} -dependent information and begin to evolve in the same way. It is this property which limits the number of momentum configurations which can be enhanced by divergences, and ultimately leads to the possibility of factorization. Mathematically, this means that we require only asymptotic information about the dependence of each mode function on k and τ .¹¹

Asymptotic behaviour. We will prove the factorization property in two parts. The first step is to show that each mode function has a ‘gap’ in its asymptotic expansion in the limit $|k\tau| \rightarrow 0$,

$$\Psi_i(k, \tau) = \psi_i(k, k_*) \left[A_i(k, k_*, \tau) + \mathcal{O}(k\tau)^2 \right]. \quad (4.8)$$

Here, $\Psi_i(k, \tau)$ is a mode function for the scale k and may depend on other labels which we collectively denote i . The prefactor $\psi_i(k, k_*)$ is time-independent and adjusted so that each Feynman propagator constructed from Ψ_i has the correct normalization. If the calculation were carried to all orders in the slow-roll expansion then Ψ_i would be independent of the reference scale k_* , but we allow ψ_i and A_i to include an explicit dependence when truncated to any finite order. In addition, $A_i(k, k_*, \tau)$ should be slowly varying, by which we mean an arbitrary polynomial in $\ln \tau$. Therefore there is a ‘gap’ in the asymptotic expansion caused by the absence of a term linear in $k\tau$. If the elementary wavefunctions satisfy (??) then so do their cosmic-time derivatives. Where required we treat these as wavefunctions of different species, distinguished by the label i .

¹¹In models containing extra characteristic scales, such as the horizon scale associated with a step or other sharp feature in the potential (see eg., Refs. [? ?]), this may no longer be true. In particular, it may happen that interactions continue outside the horizon until some characteristic time, and then switch off. In these models it need not be possible to analyse the divergence structure of the vertex integral using only asymptotic information, and extra momentum configurations may receive large enhancements. We do not consider such scenarios in this paper.

For a background spacetime sufficiently close to de Sitter, Eq. (??) can be proved using the methods of Weinberg’s theorem [?]. Weinberg constructed asymptotic approximations to the mode function of a *massless* scalar field by making an expansion in powers of k^2 . The inclusion of a slowly varying mass promotes the constants in Weinberg’s computation to arbitrary polynomials of logarithms. Because the expansion is in powers of k^2 there is automatically a gap in the asymptotic expansion of the growing mode. If the decaying mode were to begin at relative order $k\tau$ this property could be lost. However, Weinberg’s analysis shows that the decaying mode begins at relative order $(k\tau)^3$, leading to the estimate (??).

The same conclusion can be reached by studying the behaviour of solutions to the field equation in the asymptotic future $\tau \rightarrow 0$ [?]. The field equation is

$$\frac{d^2\phi}{d\tau^2} + 2aH\frac{d\phi}{d\tau} + m^2a^2\phi = 0. \quad (4.9)$$

Assuming quasi-de Sitter expansion this is an equidimensional equation with solutions $\phi \sim \tau^\Delta$, where

$$\Delta = \frac{3}{2} \pm \sqrt{\frac{9}{4} - \frac{m^2}{H^2}}. \quad (4.10)$$

In the massless case we conclude $\Delta = 0, 3$. As above, corrections due to a slowly varying mass occur as logarithms and subleading terms in the gradient expansion appear as powers of $(k/a)^2$.

4.2.1 Divergence structure of vertex integral

The second step uses (??) to show that the vertex integral produces divergences only in very restricted combinations. This point was emphasized by Weinberg [?], but here we give a more detailed analysis. It is possible to regard the theorem proved in this section as a sharper version of Weinberg’s theorem—providing control not only over power-law divergences, but also the way in which logarithmic divergences appear. On the other hand, Weinberg’s theorem applies to all orders in the loop expansion whereas our argument applies only to the bispectrum at tree-level. It would be of considerable interest to strengthen this result to constrain the possible momentum combinations which can be logarithmically enhanced at loop level.

A generic contribution to the vertex integration will mix three ‘internal’ wavefunctions, evaluated at the time η of the vertex, and three ‘external’ wavefunctions which are evaluated at the time of observation τ . The labels for the internal and external parts need not agree. For example, they may be mixed by an off-diagonal mass matrix as in (??) or (??), or the internal wavefunctions may be differentiated with respect to time. Also, depending on the assignment of $+$ and $-$ type vertices, half of the wavefunctions will be conjugated with respect to the other half. These details are irrelevant for the purposes of our discussion. We denote the external labels $\{e_1, e_2, e_3\}$ and the internal labels $\{i_1, i_2, i_3\}$. In conclusion, the late-time behaviour of each diagram can be deduced from a sum of terms of the form

$$\begin{aligned} I_n = & \left(\prod_{m \in \{1,2,3\}} \psi_{e_m}(k_m) \psi_{i_m}(k_m) \left[A_{e_m}(k_m, \tau) + \cdots \right] \right) \\ & \times \int_{-\infty}^{\tau} \frac{d\eta}{\eta^{4-2n}} f_{e_1 e_2 e_3 i_1 i_2 i_3}(\eta) \left[A_{i_1}(k_1, \eta) + \cdots \right] \left[A_{i_2}(k_2, \eta) + \cdots \right] \left[A_{i_3}(k_3, \eta) + \cdots \right], \end{aligned} \quad (4.11)$$

where $f_{e_1 e_2 e_3 i_1 i_2 i_3}$ is also taken to be slowly varying in the sense described above. For perturbatively massless fields n is a nonnegative integer arising from powers of the scale factor $a \sim -1/\tau$. We begin with four powers of a from the integration measure $\sqrt{-g}$. Further powers appear in combination with spatial gradients, which each gradient producing a power of $1/a$. Inverse gradients such as ∂^{-2} produce positive powers of a . The precise value of n depends on the *net* number of spatial gradients carried by the operator giving rise to (??).

More than two spatial gradients: $n \geq 2$. There are only two choices which produce divergences in the vertex integral: $n = 0$ and $n = 1$. We will study these in detail below. For $n \geq 2$, the vertex integral in (??) is convergent. [Clearly I_n itself may still be divergent because of $\ln \tau$ terms from the external wavefunction factors $A_{e_m}(k_m, \tau)$. Divergences of this type are not under discussion in the present section.]

4.2.2 Two spatial gradients

Now consider the case of two net spatial gradients. The primitive degree of divergence of the integral in (??) is $\int^\tau d\eta/\eta^2 \sim 1/\tau$. Therefore decaying parts of the internal wavefunctions contribute only to the convergent part of the integral, and decaying parts of the external wavefunctions produce contributions which decay at least as fast as τ .

Vertex integral. To proceed we isolate the integral in (??), which measures the cumulative effect of three-body interactions. For $n = 1$ these interactions are suppressed by a net-positive number of spatial gradients $\sim k/a$, and we should expect them to switch off in the superhorizon limit $k/a \rightarrow 0$. Hence, we anticipate that the integral produces no divergences. The late-time limit is determined by

$$\int^\tau \frac{d\eta}{\eta^2} f_{e_1 e_2 e_3 i_1 i_2 i_3}(\eta) A_{i_1}(k_1, \eta) A_{i_2}(k_2, \eta) A_{i_3}(k_3, \eta) + \text{convergent}.$$

The integral has dimensions of wavenumber or inverse time. Because f and the A_{i_m} are only slowly varying, the only scale available with the appropriate dimensions is $1/\tau$ itself. Therefore the divergent part of the integral must scale like $1/\tau$ multiplied by a polynomial in logarithms.¹² For the interactions encountered in typical inflationary theories, Weinberg's theorem guarantees that these 'fast', power-law divergences cancel [?].

Therefore, after cancellation of the 'fast' divergences, two-gradient operators with $n = 1$ behave in the same way as operators with $n \geq 2$. The vertex integral is dominated by three-body interactions which occur around the time of horizon exit, with negligible contributions from interactions occurring much later. Any divergent terms in I_1 arise only from $\ln \tau$ -dependent terms appearing in the external wavefunction factors A_{e_m} .

4.2.3 Zero spatial gradients

The remaining case is $n = 0$, generated by operators with zero net spatial gradients. In this case there is no suppression in the limit $k/a \rightarrow 0$, and therefore no expectation that these three-body interactions should switch off at late times.

¹²An alternative way to reach the same conclusion is to argue that $\int^\tau d\eta \eta^{-2} \ln^m \eta \sim \tau^{-1}$ for any power m , which can be proved by induction.

Vertex integral. As above, we first focus on the integral term in (??). There is no longer anything to prevent a ‘slow’, pure logarithmic divergence in addition to the ‘fast’ terms involving inverse powers of τ . However, we show that these slow divergences are generated by k -dependent terms drawn from a single internal wavefunction.

First, consider the contribution from the growing modes of each internal wavefunction. This can be written

$$\int^{\tau} \frac{d\eta}{\eta^4} f_{e_1 e_2 e_3 i_1 i_2 i_3}(\eta) A_{i_1}(k_1, \eta) A_{i_2}(k_2, \eta) A_{i_3}(k_3, \eta).$$

The argument of §?? shows that this diverges like τ^{-3} multiplied by a polynomial in logarithms. We will discuss these terms when we consider the external wavefunction factors. On the other hand, because of the ‘gap’ in the asymptotic expansion (??), the contribution from the $O(k\eta)^2$ -suppressed terms from two *different* wavefunctions is of the form

$$\int^{\tau} \frac{d\eta}{\eta^4} f_{e_1 e_2 e_3 i_1 i_2 i_3}(\eta) A_{i_1}(k_1, \eta) \times O(k_2 \eta)^2 \times O(k_3 \eta)^2 = \text{convergent}.$$

Therefore any pure logarithmic divergences must come from the $O(k\eta)^3$ term associated with a *single* internal wavefunction i_m . The integral has dimensions of k^3 , and hence these pure logarithms (which do not involve powers of τ) must be proportional to k_m^3 , because no other scales are available. This is the distinctive ‘local’ shape which we have been seeking.

External wavefunctions. In addition to these ‘local’ pure logarithms, the vertex integral will produce fast power-law divergences proportional to τ^{-3} and τ^{-1} . The ‘gap’ in (??) guarantees that there are no terms proportional to τ^{-2} . The τ^{-1} terms can be ignored. When multiplied by the growing modes A_{e_m} from each external wavefunction they generate contributions which are also of order τ^{-1} , and Weinberg’s theorem guarantees that they must cancel [?] ¹³. When multiplied by decaying terms in each external wavefunctions they yield contributions which decay at least as fast as τ , and therefore become negligible at late times.

The situation is different for the τ^{-3} divergences. Weinberg’s theorem likewise guarantees that, when multiplied by the growing modes from each external wavefunction, their net contribution must cancel. But it is also possible for these divergences to promote decaying terms from the external wavefunctions. As we now explain, these also produce only ‘local’ combinations of the momenta. The argument is similar to that for the internal wavefunctions.

First, consider the case where the τ^{-3} divergence combines with decaying contributions from two different external wavefunctions. The net contribution will be $\tau^{-3} \times O(\tau^2) \times O(\tau^2) = O(\tau)$. Therefore these contributions decay at late times and become negligible. It follows that pure logarithmic effects can be generated only by the $O(k\tau)^3$ term from a single external wavefunction e_m , and on dimensional grounds will be proportional to k_m^3 . There is no need to keep track of any τ^{-1} contributions generated in the same way because these are also controlled by Weinberg’s theorem.

¹³For this conclusion, we require the stronger argument of Ref. [?] which applies to operators proportional to $a^4 \sim 1/\tau^4$ in conformal time. To be clear, we would like to emphasize that Weinberg’s theorem excludes only ‘fast’ power-law divergences in correlation functions. As discussed explicitly in Ref. [?], it does not restrict the appearance of ‘slow’ logarithmic effects.

4.2.4 Inverse spatial gradients

We can not yet conclude that the vertex integration produces only divergences proportional to a pure power k_i^3 of one of the external momenta, because in Einstein gravity some modes of the metric are not dynamical: instead, they are removed by constraints. The process of solving these constraints can produce inverse spatial gradients ∂^{-2} [? ?]. Working in ADM variables, the metric can be written

$$ds^2 = -N^2 dt^2 + h_{ij}(dx^i + N^i dt)(dx^j + N^j dt), \quad (4.12)$$

where N is the lapse function and N^i is the shift vector. Only the three-metric h_{ij} carries independent degrees of freedom. The lapse and shift are determined by constraints. (See Appendix ??.) To obtain the third-order action it is only necessary to solve these constraints to first order [? ?]. We set $N_i = \partial_i \vartheta_1 + \beta_{1i}$, where β_{1i} is divergenceless. It can be ignored for the purposes of the three-point function because it does not contribute to the third-order action. The solution for ϑ_1 is [?]

$$\frac{4H}{a^2} \partial^2 \vartheta_1 = -2\dot{\phi}^\alpha \delta\phi_\alpha + \left(2\ddot{\phi}^\alpha + \frac{\dot{\phi}^2}{H} \dot{\phi}^\alpha\right) \delta\phi_\alpha. \quad (4.13)$$

In principle the shift vector $N_i = \partial_i \vartheta_1$ can appear in operators with zero net spatial gradients, generating enhanced shapes formed from rational functions of the external momenta. (If present, these ratios of the external momenta would appear as a prefactor in Eq. (??) which we have not written explicitly.) Although this outcome is compatible with the conclusions of §§??–?? these shapes are not local and could not be reproduced by separate-universe type formulae. To demonstrate that only form-factors compatible with the separate universe method are required we must show that these enhanced non-local shapes are absent.

Whether ϑ_1 can produce divergences in the vertex integral depends on its asymptotic behaviour. This is an important issue beyond the renormalization-group approach we are developing in this paper, because the decay rate of the shift vector is important in any attempt to justify the separate universe method. Weinberg gave an argument using a broken-symmetry approach [? ?]. More recently, Sugiyama, Komatsu & Futamase showed that the Einstein equations require the shift vector to decay at late times during inflation [?]. Up to next-order, it can be verified by explicit calculation that (??) gives

$$\frac{4H}{a^2} \partial^2 \vartheta_1 = \mathcal{O}(k\eta)^2. \quad (4.14)$$

This gives $\vartheta_1 = \mathcal{O}(1)$. The general analysis of Sugiyama et al. extends this conclusion to all orders in the slow-roll expansion.

Absence of time-dependent terms generated by the shift. This decay rate is not sufficiently rapid to prevent the generation of time-dependent terms from arbitrary operators involving the shift vector N_i . However, because the decay is exponentially fast in cosmic time, it will prevent divergences in any operator which involves two or more powers of $\partial^2 \vartheta_1 / a^2$. Only operators linear in $\partial^2 \vartheta_1 / a^2$ can generate time dependence.¹⁴ It can be shown that the

¹⁴In principle, terms such as $N_i N^i = \partial_i \vartheta_1 \partial_i \vartheta_1 / a^2$ can generate time dependence, even though they are quadratic in ϑ_1 . However, such contractions are not generated in Einstein gravity combined with the scalar field theories we are considering.

second-order action is entirely independent of ϑ , and that the third-order action contains no terms linear in $\partial^2 \vartheta_1/a^2$. Therefore the lapse function gives rise to no time-dependent terms in the two- or three-point functions, and no nonlocal momentum configurations can be enhanced by divergences.

4.2.5 Factorization for the three-point function

In conclusion, to all orders in the slow-roll expansion, each diagram contributing to the three-point function with external wavefunctions carrying labels $\{e_1, e_2, e_3\}$ will schematically be of the form

$$I_{e_1 e_2 e_3} = \left[\prod_m \psi_{e_m}(k_m) \psi_{i_m}(k_m) \right] A_{e_1}(k_1, \tau) A_{e_2}(k_2, \tau) A_{e_3}(k_3, \tau) \\ \times \left[u(\tau) k_1^3 + v(\tau) k_2^3 + w(\tau) k_3^3 + \text{finite} \right]_{e_1 e_2 e_3 i_1 i_2 i_3}, \quad (4.15)$$

where u , v and w are slowly-varying functions of τ (and perhaps also k or k_*) given by arbitrary polynomials of logarithms. The ‘finite’ term is independent of τ . The precise form of u , v , w and the finite piece depends on the interaction under discussion. The notation $[\cdots]_{e_1 e_2 e_3 i_1 i_2 i_3}$ indicates that both the divergent and finite pieces produced by the vertex integral can depend on the labels for the internal and external wavefunctions. Eq. (??) is the formal statement of the factorization theorem.¹⁵

As for the two-point function, perturbative calculations using quantum field theory will not provide us with the functional form of u , v , or w . Instead, they generate terms in the Taylor series expansion for these functions based at the arbitrary horizon-crossing time for k_* . The full functions must be reverse-engineered using the methods of the renormalization group.

Eq. (??) clearly exhibits the Sudakov-like enhancements generated by divergent logarithms. In agreement with the discussion of §??, we see that large Sudakov-like effects in each external wavefunction can be absorbed into the form factor $\Gamma_{\alpha i}$. This accounts for the first line in (??). Also, we now see that (??) guarantees the vertex integral will produce precisely the structure identified in §?? as a prerequisite for the second line in (??).

To develop a renormalization-group approach to the three-point function we will not pursue the factorization of this form-factor directly. Instead we follow the approach of §??. Specializing to a multiple-field model in which the external labels $\{e_1, e_2, e_3\}$ correspond to the different species of scalar fields, we see that the lowest terms in the Taylor expansion

¹⁵As described in §??, the time-dependent ‘local’ contributions u , v and w may include terms generated by promotion of decaying terms in an external wavefunction Ψ_{e_m} due to τ^{-3} divergences in the vertex integral. In (??) we have redefined these contributions by dividing out a factor of the corresponding external growing mode A_{e_m} . This is harmless overall, leading only to a redefinition of the associated source term. No such terms have yet been encountered in practical calculations, so there has been no need to keep track of this redefinition. The first term of this type would arise from subleading corrections to the $V''' \delta\phi^3$ vertex, which already contributes at next-order. Therefore the effect is *next*-next-order in fixed-order perturbation theory and would contribute only to the next-leading-logarithm solution of the renormalization group equation. (See footnote ?? on p. ??.)

around the horizon-crossing time for k_* are

$$\begin{aligned}
I_{\alpha\beta\gamma} \approx & I_{\alpha\beta\gamma}^* - [u_{\alpha\lambda}^* I_{\lambda\beta\gamma}^* + u_{\beta\lambda}^* I_{\alpha\lambda\gamma}^* + u_{\gamma\lambda}^* I_{\alpha\beta\lambda}^*] \ln(-k_*\tau) \\
& + \left[\prod_m |\psi_\alpha(k_m)|^2 \right] A_{\alpha\lambda}^*(k_1) A_{\beta\mu}^*(k_2) A_{\gamma\nu}^*(k_3) \left. \frac{du_{\lambda\mu\nu}}{dN} \right|_* k_1^3 \ln(-k_*\tau) \\
& + \text{permutations,}
\end{aligned} \tag{4.16}$$

where $du_{\lambda\mu\nu}/dN$ is defined by this expression, and ‘permutations’ includes the terms generated from the Taylor expansion of $v(\tau)k_2^3$ and $w(\tau)k_3^3$ in (??). The Taylor coefficient $u_{\alpha\beta}$ associated with time-dependence of the external wavefunction factors is already known. Therefore to reverse-engineer the full functional form of $I_{\alpha\beta\gamma}$ we require only an estimate of the derivative $du_{\lambda\mu\nu}/dN|_*$, which can be read off from the $\ln(-k_*\tau)$ terms in a next-order computation of the three-point function. We collect the details of this calculation in Appendix ??.

4.3 Renormalization group analysis for the three-point function

We can now use (??) to apply the dynamical renormalization group to the three-point function. The procedure is very similar to the two-point function analysis in §??, although potentially complicated because the three-point function depends on three distinct external momenta k_1, k_2, k_3 . In addition to a hierarchy between the scale of these momenta and the horizon scale aH , there may now be hierarchies among the external momenta themselves.

Shape and scale effects. The three-point function enforces momentum conservation $\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3 = 0$ for the wavevectors \mathbf{k}_i carried by its external lines. Therefore the \mathbf{k}_i can be regarded as forming a triangle.

In addition to the time-dependent logarithms $\ln|k_*\tau|$ which we have studied in §§??–??, it is now possible to generate logarithms involving ratios of the form $\ln k_i/k_t$ and $\ln k_t/k_*$, where $k_t = k_1 + k_2 + k_3$ is the perimeter of the triangle. (Logarithms such as $\ln k_i/k_*$ can be converted to $\ln k_i/k_t + \ln k_t/k_*$.) These effects were discussed by Burrage et al. [? ?], who identified them with a response to changes in the size or shape of the momentum triangle. When any of these hierarchies become large, resummation of the corresponding logarithms will endow the three-point function with a new, enhanced Sudakov-like structure.

We can regard k_t as an analogue of the single scale k in (??). Indeed, the term $\ln 2k/k_*$ appearing in (??) can be regarded as $\ln k_t/k_*$, with $k_t = k_1 + k_2$ the ‘perimeter’ of the momentum 2-gon. If the momentum k_i associated with each side of the triangle is not too different from k_t then there is only one relevant momentum scale and situation is very similar to that of the two-point function. This is the ‘equilateral’ configuration. The converse situation occurs if one momentum is very much smaller than k_t . In this case there are multiple hierarchies and computation of the three-point function becomes more complex. In this paper we focus only on the case where $k_i \sim k_t$ and there is a single hierarchy. We intend to return to the question of multiple hierarchies in a future publication.

External wavefunctions. We divide the analysis into time-dependent terms arising from external wavefunction factors, and those arising from the vertex integral. The time-dependence associated with external wavefunction factors is ‘unsourced’, in the sense that new contributions are not continuously generated. (One way to regard the time-dependent terms generated by external wavefunctions is as a resummation of the infinite sequence of Feynman diagrams

generated by arbitrary insertions of the mass operator $m_{\alpha\beta}\delta\phi_\alpha\delta\phi_\beta$ on each external line.) In contrast, time-dependence arising from the vertex integral is actively ‘sourced’ by ongoing three-body interactions, as described in §??.

The factorization theorem (??) shows how to deal with these unsourced terms. We define a ‘bispectrum’ $B_{\alpha\beta\gamma}$ by

$$\langle\delta\phi_\alpha(\mathbf{k}_1)\delta\phi_\beta(\mathbf{k}_2)\delta\phi_\gamma(\mathbf{k}_3)\rangle = (2\pi)^3\delta(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3)\frac{B_{\alpha\beta\gamma}}{4\prod_i k_i^3}. \quad (4.17)$$

The factor $(4\prod_i k_i^3)^{-1}$ is generated from products of the ψ_i which normalize each wavefunction in (??). The factorization theorem shows that the ‘sourced’ and ‘unsourced’ pieces contribute additively to $B_{\alpha\beta\gamma}$ and can therefore be considered separately. Therefore we ignore all ‘sourced’ contributions from the vertex integral, which will be dealt with below.

Excluding the remaining soft effects generated by powers of $\ln|k_*\tau|$ from each external wavefunction, we obtain a ‘hard’ component $B_{\alpha\beta\gamma}^{\text{hard}*}$ analogous to (??). Because there is no hierarchy between the external momenta, the hard component can be computed using fixed-order perturbation theory. Its lowest-order contribution was calculated in Ref. [?]. Alternatively it may be obtained by extracting the lowest-order terms from Eqs. (??), (??) and (??). This contribution is analogous to the lowest-order term $H_*^2\delta_{\alpha\beta}$ in (??). In Appendix ?? we extend the results of Ref. [?] by computing the three-point function to next-order. The subleading contributions from this calculation which are not enhanced by soft logarithms are analogous to the next-order term $r_{\alpha\beta}^*$ in (??). However, for this discussion in this section we will not need an explicit expression for $B_{\alpha\beta\gamma}^{\text{hard}*}$.

Comparison with (??) shows that the unsourced contribution to the bispectrum, which we denote $\hat{B}_{\alpha\beta\gamma}$, has the Taylor series expansion

$$\hat{B}_{\alpha\beta\gamma} \simeq B_{\alpha\beta\gamma}^{\text{hard}*} - \left[u_{\alpha\lambda}^* B_{\lambda\beta\gamma}^{\text{hard}*} + u_{\beta\lambda}^* B_{\alpha\lambda\gamma}^{\text{hard}*} + u_{\gamma\lambda}^* B_{\alpha\beta\lambda}^{\text{hard}*} \right] \ln(-k_*\tau) + \dots \quad (4.18)$$

We now apply the argument of §?? to obtain the dynamical renormalization-group equation

$$-\frac{d\hat{B}_{\alpha\beta\gamma}}{d\ln\tau} = u_{\alpha\lambda}\hat{B}_{\lambda\beta\gamma} + u_{\beta\lambda}\hat{B}_{\alpha\lambda\gamma} + u_{\gamma\lambda}\hat{B}_{\alpha\beta\lambda}. \quad (4.19)$$

A suitable initial condition can be determined from the fixed-order perturbative expression for $B_{\alpha\beta\gamma}^{\text{hard}*}$. As for the two-point function, the methods of Ref. [?] can be used to show that (??) is equivalent to absorption of the time-dependent logarithms in the form-factor $\Gamma_{\alpha i}$.

Vertex integral. We now return to the ‘sourced’ contributions generated by the vertex integral which we introduced in Eq. (??). Collecting terms from Appendix ??, we find

$$B_{\alpha\beta\gamma}^{\text{div,vertex}} \supseteq -H_*^4 u_{\alpha\beta\gamma}^* \ln(-k_*\tau) \sum_i k_i^3, \quad (4.20)$$

where a sub- or superscript ‘*’ denotes evaluation at the time $|k_*\tau| = 1$, and $u_{\alpha\beta\gamma}$ is obtained by differentiating the expansion tensor $u_{\alpha\beta}$, given in Eq. (??),

$$u_{\alpha\beta\gamma} = \partial_\gamma u_{\alpha\beta} = -\frac{V_{\alpha\beta\gamma}}{3H^2} + \frac{\dot{\phi}_\alpha}{H} u_{\beta\gamma} + \frac{\dot{\phi}_\beta}{H} u_{\alpha\gamma} + \frac{\dot{\phi}_\gamma}{H} u_{\alpha\beta} - \frac{\dot{\phi}_\alpha \dot{\phi}_\beta \dot{\phi}_\gamma}{H^3}. \quad (4.21)$$

It is symmetric under exchange of any two indices. The superscript ‘div,vertex’ is a reminder that Eq. (??) includes only divergent terms from the vertex integral. We see that time-dependent logarithms multiply only the local shapes k_1^3 , k_2^3 and k_3^3 , as required by the factorization theorem. The symmetries of $u_{\alpha\beta\gamma}$ reduce these to the combination $\sum_i k_i^3$.

To apply a renormalization-group analysis, we introduce a form-factor for each shape which can be logarithmically enhanced. In this case these are the local combinations,

$$B_{\alpha\beta\gamma}^{\text{sourced}} = a_{\alpha|\beta\gamma} k_1^3 + a_{\beta|\alpha\gamma} k_2^3 + a_{\gamma|\alpha\beta} k_3^3. \quad (4.22)$$

The form-factors $a_{\alpha|\beta\gamma}$ are symmetric under exchange of β and γ , but need have no other symmetries. The full bispectrum is obtained by adding these sourced contributions to the unsourced terms obtained from (??).

As usual, exclusion of all terms enhanced by soft logarithms yields ‘hard’ components for the $a_{\alpha|\beta\gamma}$, and in the absence of other large hierarchies they can also be computed using fixed-order perturbation theory. Comparison of (??), (??) and (??) shows that the $a_{\alpha|\beta\gamma}$ have a Taylor series expansion

$$a_{\alpha|\beta\gamma} \simeq a_{\alpha|\beta\gamma}^{\text{hard*}} - \left[u_{\alpha\lambda}^* a_{\lambda|\beta\gamma}^{\text{hard*}} + u_{\beta\lambda}^* a_{\alpha|\lambda\gamma}^{\text{hard*}} + u_{\gamma\lambda}^* a_{\alpha|\beta\lambda}^{\text{hard*}} + u_{\alpha\lambda\mu}^* \Sigma_{\lambda\beta}^{\text{hard*}} \Sigma_{\mu\gamma}^{\text{hard*}} \right] \ln(-k_*\tau) + \dots \quad (4.23)$$

We have replaced H_*^4 by a suitable product of $\Sigma_{\alpha\beta}^{\text{hard*}}$ which preserves the index symmetries. As in §?? this may generate a small mismatch at next-next-order, which translates to an error in the solution of the renormalization group which is below leading-logarithm accuracy. The renormalization-group prescription gives

$$-\frac{da_{\alpha|\beta\gamma}}{d\ln\tau} = u_{\alpha\lambda} a_{\lambda|\beta\gamma} + u_{\beta\lambda} a_{\alpha|\lambda\gamma} + u_{\gamma\lambda} a_{\alpha|\beta\lambda} + u_{\alpha\lambda\mu} \Sigma_{\lambda\beta} \Sigma_{\mu\gamma}. \quad (4.24)$$

This is the transport equation for $a_{\alpha|\beta\gamma}$ derived in Ref. [?], where it was demonstrated that its solution generates the form-factor $\partial^2\phi_\alpha/\partial\phi_i^*\partial\phi_j^*$. Combining the sourced and unsourced components, we reproduce the anticipated expression (??).

There is some arbitrariness in choosing an initial condition for $a_{\alpha|\beta\gamma}$. We are free to treat any local terms which appear in the lowest-order correlation function $B_{\alpha\beta\gamma}^{\text{hard*}}$ *either* as contributions to the unsourced part $\hat{B}_{\alpha\beta\gamma}$, *or* an initial condition for $a_{\alpha|\beta\gamma}$.¹⁶ Whichever choice we make the result is the same, because the results of Ref. [?] show that this initial condition contributes additively to the final bispectrum.

5 Applications

In this section we briefly discuss two examples which illustrate the utility of the renormalization group framework.

5.1 Matching between quantum and classical eras

In §§??–?? we emphasized the separation between ‘hard’, model-independent processes by which inflationary fluctuations are created, and ‘soft’, model-dependent processes by which

¹⁶For clarity, we note that this is not the same ambiguity discussed in footnote ?? on p. ??, which was resolved by a choice of factorization scheme.

they evolve. The hard creation process is associated with loss of a decaying mode, causing interference effects to cease and the fluctuations to behave classically. Usually, the separate universe picture is considered as a framework which can be used to evolve these classicalized fluctuations [? ? ?].

Power-law corrections. Implicit in this point of view is a matching between the quantum and classical parts of the calculation. Beginning with Polarski & Starobinsky [? ?], several authors have noticed that this leads to a potential ambiguity. The issue was later studied in more detail by Leach & Liddle [?] and has recently been revisited by Nalson et al. [?].

The problem can be stated simply. In Eqs. (??) and (??), decaying power-law corrections which scale like positive powers of $|k\tau|$ have been neglected. These already occur at lowest order in slow roll, where the equal-time two-point function at an arbitrary conformal time τ can be written

$$\langle \delta\phi_\alpha(\mathbf{k}_1)\delta\phi_\beta(\mathbf{k}_2) \rangle_\tau = (2\pi)^3 \delta(\mathbf{k}_1 + \mathbf{k}_2) \frac{H_*^2}{2k^3} (1 + k^2\tau^2). \quad (5.1)$$

The term $k^2\tau^2$ is unity at horizon crossing but decays exponentially fast outside the horizon, leading to evolution of the typical fluctuation amplitude $\langle \phi^2 \rangle^{1/2}$ by a factor of $\sqrt{2}$ between horizon crossing and a few e-folds later. By itself this is not significant because one should not think of a measurable classical fluctuation with fixed amplitude until the decaying mode is lost, which corresponds to $k^2\tau^2$ becoming negligible [?]. However, it does demonstrate that (from this point of view) classical reasoning can not be used until at least a few e-folds outside the horizon. A similar issue will exist for higher n -point functions.

Choice of matching surface. If the separate universe method is regarded in classical terms we cannot apply it at the moment of horizon exit. But Eqs. (??) and (??) also show that we cannot wait too long before switching to the classical calculation, otherwise this quantum initial condition will be invalidated by the growing logarithm $\ln(-k_*\tau)$. At the earliest, we could perhaps consider the fluctuations to be approximately classical when they are ~ 2 e-folds outside the horizon, making $k^2\tau^2 \sim e^{-4} \approx 0.018$. In Fig. ?? we show the effect of different matching prescriptions in the double-quadratic model (??), using the k -scale, initial conditions and model parameters of Fig. ?. We plot the observable quantity P_ζ , which can be obtained from $\Sigma_{\alpha\beta}$ by a gauge transformation [?]. We compute $\Sigma_{\alpha\beta}$ numerically using Eq. (??) and making use of the slow-roll approximation.

In Fig. ?? we perform the matching at $N = 2$, $N = 4$ and $N = 6$ e-folds after horizon crossing. Initial conditions for the classical evolution are set using the lowest-order approximation with $\Sigma_{\alpha\beta} = H_N^2 \delta_{\alpha\beta}$, where H_N is the value of the Hubble rate on the matching surface. The decaying power-law term $k^2\tau^2$ is ignored in setting these initial conditions.¹⁷

In Fig. ?? the matching is also performed at $N = 2$, $N = 4$ and $N = 6$ e-folds after horizon crossing, but with the initial condition $\Sigma_{\alpha\beta} = H_k^2 \delta_{\alpha\beta}$ where H_k is the Hubble rate at the moment of horizon exit. Finally, in Fig. ?? we use the full NLO expression (??) to provide

¹⁷In §?? we argued that the decaying mode began at $O(k\tau)^3$, and therefore this term must originate from the growing mode which dominates the subsequent classical evolution. Hence, its contribution could be included in a self-consistent calculation. This possibility was suggested by Polarski & Starobinsky [?] and Lalak et al. [?].

We have explained in §?? that decaying terms of the form $k^2\tau^2$ do not contribute the DRGE evolution, and for this reason we have not retained it when setting initial conditions.

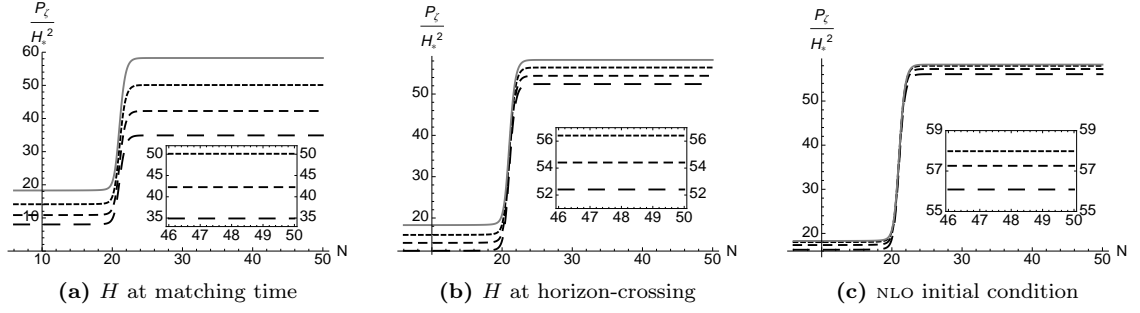


Figure 5: Impact of different matching prescriptions on the final value of the power spectrum of ζ in the double-quadratic model (??). The quoted value is for the scale which passed outside the horizon with field values $\varphi = 8.2M_P$ and $\chi = 12.9M_P$, labelled $N = 0$ on the figures. In each figure, the short dashed line corresponds to beginning the classical calculation at $N = 2$ e-folds after horizon exit; the medium dashed line at $N = 4$ e-folds; and the long dashed line at $N = 6$ e-folds. The evolution generated by the dynamical renormalization group is the grey solid line. In Fig. ??, initial conditions are chosen using $\Sigma_{\alpha\beta} = H_N^2 \delta_{\alpha\beta}$, with H_N the value of the Hubble rate at the matching time. In Fig. ?? the matching is performed using $\Sigma_{\alpha\beta} = H_k^2 \delta_{\alpha\beta}$, where H_k is the value of the Hubble rate at horizon crossing. In Fig. ?? the full NLO initial condition (??) is used, setting $k_* = k$ to minimize the $\ln k/k_*$ terms. In each plot the inset panel shows a close-up of the last four e-folds of evolution.

an initial condition at each matching surface. In each figure, the solid grey line shows the evolution computed using the dynamical renormalization group (see below).

With the full NLO initial condition (Fig. ??) the discrepancy between the asymptotic value of P_ζ computed using each prescription, and also the DRGE evolution, is a few percent. In Fig. ?? the discrepancy between different matching prescriptions is of order 3%–7%. The discrepancy between the $N = 6$ prescription and the DRGE evolution is as large as 10%. In Fig. ?? the discrepancies are very large, between 10% and 40%.

In Fig. ??, the expression (??) is used to provide an estimate of $\Sigma_{\alpha\beta}$ at the matching surface. One might have expected this estimate to be accurate because it is computed to next-order, and the slow-roll parameters are individually small. Therefore we might also have expected the final difference in P_ζ to be negligible, because the separate universe method is supposed to be independent of the time at which initial conditions are set. The discrepancy arises because, with matching performed at $N > 0$, the fixed-order expression is used to evolve $\Sigma_{\alpha\beta}$ up to N e-folds. With the matching performed at $N = 0$, all-orders resummation is used. Hence, the non-negligible difference between each line in Fig. ?? measures the discrepancy between the resummed and fixed-order calculations. It would be even more significant if these first N e-folds coincided with stronger evolution, such as that occurring near the spike in Fig. ??.

DRGE analysis. As observations improve, an ambiguity of order 10% will certainly be important. Therefore we would like to understand which of these prescriptions, if any, is correct. This question cannot be answered within a classical framework. The resolution is that there is no requirement to divide the calculation into ‘quantum’ and ‘classical’ eras joined by a matching condition. Although this is a useful picture which guides our thinking, the calculation can be carried out entirely within the framework of quantum field theory and yields a unique answer. Indeed, as we have seen in §§??–??, the relevant division is not into ‘quantum’ and ‘classical’ effects, but ‘hard’ and ‘soft’ processes. It simply happens that,

because interference effects are absent in the soft superhorizon evolution, it can be described at tree-level using the classical equations of motion. In the quantum field theory calculation there is no arbitrary matching between ‘quantum’ and ‘classical’ eras, but only the floating factorization scale. Changes in the factorization scale leave all correlation functions invariant.

The renormalization-group approach we have described allows each correlation function to be calculated purely using the methods of quantum field theory and therefore provides a unique resolution of this matching ambiguity. The prescription can be extracted from Eqs. (??) and (??). Positive powers of $k\tau$ harmlessly decay and do not contribute to the Taylor series in $\ln(-k_*\tau)$ at late times. When reversing-engineering this Taylor expansion, the renormalization group evolution should begin at the expansion point k_* and the initial condition should be the constant term in the Taylor series. As we argued in §??, to obtain the most accurate answer we should evaluate (??) at the horizon-crossing time for k . Therefore we should simultaneously begin the renormalization-group evolution at this time. Hence the prescription provided by the DRGE is to use the initial condition $\Sigma_{\alpha\beta} \approx H_k^2 \delta_{\alpha\beta}$ but to begin the ‘classical’ evolution at $N = 0$. If desired, higher-order corrections can be retained to give a more accurate estimate of the initial value. This has been done to obtain the grey lines in Fig. ??, although the change due to inclusion of NLO terms in the DRGE initial condition is only of order 0.3%.

The prescription provided by the dynamical renormalization group is incompatible with classical intuition, but by itself this is not important because no classical phase of evolution is being invoked in this calculation.

5.2 Nontrivial field-space metric

Elliston et al. [?] recently calculated the inflationary bispectrum generated by a σ -model Lagrangian with nontrivial field-space metric. The action is

$$S = \frac{1}{2} \int d^3x dt a^3 \left[M_{\text{P}}^2 R - \mathbf{G}_{\alpha\beta} \partial_a \phi^\alpha \partial^a \phi^\beta - 2V \right] \quad (5.2)$$

where V is a potential and $\mathbf{G}_{\alpha\beta}$ is an arbitrary symmetric matrix which can be interpreted as a metric.

To preserve manifest covariance with respect to $\mathbf{G}_{\alpha\beta}$ it is helpful to describe fluctuations using a variable which obeys a covariant transformation law [? ?]. Gong & Tanaka gave a prescription for rewriting inflationary perturbation theory in terms of such a variable, which they denoted Q^α , and computed the action for field fluctuations to third-order [?].

Two-point function. The equal-time two-point function for Q^α had already been obtained to leading order by Sasaki & Stewart [?]. The analogue of (??) was computed by Nakamura & Stewart [?]. There is a correction to the mass matrix and a more complex structure due to preservation of manifest covariance,¹⁸

$$\begin{aligned} \langle Q^\alpha(\mathbf{k}_1) Q^\beta(\mathbf{k}_2) \rangle &\supseteq (2\pi)^3 \delta(\mathbf{k}_1 + \mathbf{k}_2) \frac{H_*^2}{2k_*^3} \mathbf{\Pi}^\alpha_i \mathbf{\Pi}^\beta_j \\ &\times \left\{ \mathbf{G}^{ij} \left[1 + 2\varepsilon_* \left(1 - \gamma_E - \ln \frac{2k}{k_*} \right) \right] + 2\mathbf{w}^{ij} \left[2 - \ln(-k_*\tau) - \ln \frac{2k}{k_*} - \gamma_E \right] \right\}. \end{aligned} \quad (5.3)$$

¹⁸In Ref. [?] the parallel propagator matrices $\mathbf{\Pi}^\alpha_i$ were omitted. Because the metric is determined by parallel transport, $\mathbf{\Pi}^\alpha_i \mathbf{\Pi}^\beta_j \mathbf{G}^{ij} = \mathbf{G}^{\alpha\beta}$. Therefore, provided we carefully keep track of the evaluation point for the metric, these matrices are not needed explicitly at lowest-order.

In accordance with our usual convention we have neglected decaying power-law corrections. Greek indices α, β, \dots label the tangent space at field-space coordinate $\phi^\alpha(\tau)$, whereas indices i, j, \dots label the tangent space at coordinate $\phi^\alpha(\tau_*)$, where τ_* is the horizon-crossing time for a reference scale k_* . The expansion tensor \mathbf{w}_{ij} satisfies

$$\mathbf{w}_{ij} = -\frac{V_{ij}}{3H^2} + \frac{1}{3H^2} \frac{1}{a^3} \frac{D}{dt} \left(\frac{a^3}{H} \dot{\phi}_i \dot{\phi}_j \right) + \frac{1}{3} \mathbf{R}_{imnj} \frac{\dot{\phi}^m}{H} \frac{\dot{\phi}^n}{H}. \quad (5.4)$$

The indices i, j imply that (??) is evaluated at time τ_* . The same applies for \mathbf{G}_{ij} . The derivative D/dt is $\dot{\phi}^\alpha \nabla_\alpha$, where ∇_α is the covariant derivative compatible with $\mathbf{G}_{\alpha\beta}$ and $\mathbf{R}_{\alpha\beta\gamma\delta}$ is the corresponding Riemann tensor. Finally, the parallel propagator $\mathbf{\Pi}^\alpha_i$ expresses parallel transport along the inflationary trajectory in field-space between $\phi^\alpha(\tau_*)$ and $\phi^\alpha(\tau)$,

$$\mathbf{\Pi}^\alpha_i = \mathcal{T} \exp \left(- \int_{\tau_*}^{\tau} d\eta \Gamma_{\beta'\gamma'}^{\alpha'} d\phi^{\beta'} \right) \delta_i^{\gamma'}. \quad (5.5)$$

Indices α', β', \dots label the tangent space associated with the integration time η , and \mathcal{T} is the time-ordering operator. For all details we refer to Elliston et al. [?].

Three-point function. The equal-time three-point function for Q^α was obtained up to $O(\dot{\phi}/H)^2$ in Ref. [?]. It is

$$\langle Q^\alpha(\mathbf{k}_1) Q^\beta(\mathbf{k}_2) Q^\gamma(\mathbf{k}_3) \rangle \supseteq (2\pi)^3 \delta(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3) \frac{H_*^4}{4 \prod_i k_i^4} \mathbf{\Pi}^\alpha_i \mathbf{\Pi}^\beta_j \mathbf{\Pi}^\gamma_k A^{ijk}(N), \quad (5.6)$$

where A^{ijk} transforms as a three-tensor in the tangent space at time τ_* and is defined by

$$\begin{aligned} A^{ijk} = & \frac{1}{M_{\text{Pl}}^2} \frac{\dot{\phi}^i}{H} \mathbf{G}^{jk} \left(\frac{k_1}{2} \mathbf{k}_2 \cdot \mathbf{k}_3 - 2 \frac{k_2^2 k_3^2}{k_t} \right) \\ & + \frac{4}{3} \mathbf{R}^{i(jk)m} \frac{\dot{\phi}_m}{H} \left[k_1^3 \left(\gamma_E + \ln(-k_* \tau) + \ln \frac{k_t}{k_*} \right) - k_1^2 k_t + \frac{k_1^2 k_2 k_3}{k_t} \right] \\ & + \frac{1}{3} \nabla^{i(jk)m} \mathbf{R}^{j|mn|k} \frac{\dot{\phi}_m}{H} \frac{\dot{\phi}_n}{H} \left[-k_1^3 \left(\ln(-k_* \tau) + \ln \frac{k_t}{k_*} + \gamma_E + \frac{1}{3} \right) + \frac{4}{9} k_t^3 - k_t K^2 \right] \\ & - \frac{4}{3} \nabla^n \mathbf{R}^{i(jk)m} \frac{\dot{\phi}_m}{H} \frac{\dot{\phi}_n}{H} \left[\frac{k_1^3}{2} \left(\ln^2(-k_* \tau) - \gamma_E^2 + \frac{\pi^2}{12} - \left[2\gamma_E + \ln \frac{k_t}{k_*} \right] \ln \frac{k_t}{k_*} \right) \right. \\ & \quad \left. + k_1^2 k_t \left(\ln \frac{k_t}{k_*} + \gamma_E - 1 \right) - \frac{k_1^2 k_2 k_3}{k_t} \left(\gamma_E + \ln \frac{k_t}{k_*} \right) \right] \\ & + \text{cyclic}. \end{aligned} \quad (5.7)$$

We have used $k_t = k_1 + k_2 + k_3$ and $K^2 = \sum_{i < j} k_i k_j = k_1 k_2 + k_1 k_3 + k_2 k_3$. As in (??), all time-dependent quantities are evaluated at the horizon-crossing time for k_* .

Eqs. (??)–(??) are considerably more complicated than their flat field-space counterparts.¹⁹ Growing terms involving $\ln(-k_* \tau)$ are already present at lowest nontrivial order in $\dot{\phi}/H$. At $O(\dot{\phi}/H)^2$ a double-logarithmic term involving $\ln^2(-k_* \tau)$ appears. In flat field-space, growing terms appear only at $O(\dot{\phi}/H)^3$ and double logarithms do not occur until even higher orders. In combination with the propagator terms $\mathbf{\Pi}^\alpha_i$ these make the time dependence of (??) appear extremely complicated.

¹⁹The first line in (??) covariantizes the lowest-order flat-space computation given in Ref. [?]. If next-order corrections were to be kept they would covariantize the computation described in Appendix ??.

Dynamical renormalization group analysis. Eqs. (??), (??) and (??) present a case study in the use of renormalization group techniques to extract a Sudakov-like structure from logarithmic divergences.

The Taylor theorem on a curved manifold is

$$A^{\alpha\cdots}_{\beta\cdots}(x) = \mathbf{\Pi}^{\alpha}_{\alpha'}(x, x') \cdots \mathbf{\Pi}^{\beta'}_{\beta}(x, x') \cdots \left[A^{\alpha'\cdots}_{\beta'\cdots} \Big|_{x'} + \frac{D}{d\lambda} A^{\alpha'\cdots}_{\beta'\cdots} \Big|_{x'} \delta\lambda + \cdots \right], \quad (5.8)$$

where the omitted terms are $O(\delta\lambda^2)$. To write (??) we have assumed a smooth path connecting x and x' , with $\mathbf{\Pi}^{\alpha}_{\alpha'}(x, x')$ the parallel propagator along this path and $D/d\lambda$ the corresponding parametric derivative. As in §??, the Taylor theorem (??) is our principal tool in the renormalization-group analysis. The path connecting x and x' will be the inflationary trajectory, and the parametric derivative along it will be D/dN .

The argument now parallels §?? precisely. Comparing Eq. (??) with (??) allows us to conclude

$$\frac{D}{dN} \Sigma^{\alpha\beta} \Big|_{-k_*\tau=1} = 2H^2 \mathbf{w}^{\alpha\beta} \Big|_{-k_*\tau=1}. \quad (5.9)$$

The arbitrariness of k_* allows this to be promoted to a differential equation. Replacing H_* by suitable copies of $\Sigma_{\alpha\beta}$, as in §??, we find

$$\frac{D\Sigma^{\alpha\beta}}{dN} = \mathbf{w}^{\alpha}_{\gamma} \Sigma^{\gamma\beta} + \mathbf{w}^{\beta}_{\gamma} \Sigma^{\gamma\alpha}. \quad (5.10)$$

A similar analysis can be performed for the three-point function. Eq. (??) shows that all time-dependent logarithms multiply only local momentum shapes, in agreement with the factorization theorem of §??. (This is not altered by the presence of a nontrivial metric.) Repeating the argument of §??, we conclude

$$\frac{Da_{\alpha|\beta\gamma}}{dN} = \mathbf{w}^{\lambda}_{\alpha} a_{\lambda|\beta\gamma} + \mathbf{w}^{\lambda}_{\beta} a_{\alpha|\lambda\gamma} + \mathbf{w}^{\lambda}_{\gamma} a_{\alpha|\beta\lambda} + \mathbf{w}^{\lambda\mu}_{\alpha} \Sigma_{\lambda\beta} \Sigma_{\mu\gamma}, \quad (5.11)$$

where $\mathbf{w}_{\alpha\beta\gamma}$ satisfies

$$\mathbf{w}_{\alpha\beta\gamma} = \nabla_{(\alpha} \mathbf{w}_{\beta\gamma)} + \frac{1}{3} \left(\nabla_{(\alpha} \mathbf{R}_{\beta|\lambda\mu|\gamma)} \frac{\dot{\phi}^{\lambda}}{H} \frac{\dot{\phi}^{\mu}}{H} - 4\mathbf{R}_{\alpha(\beta\gamma)\lambda} \frac{\dot{\phi}^{\lambda}}{H} \right). \quad (5.12)$$

Jacobi equation. Elliston et al. deduced equations equivalent to (??) and (??)–(??) from the evolution of a covariant connecting vector between two adjacent inflationary trajectories [?], and demonstrated that (when solved perturbatively) these transport equations reproduce the ‘divergences’ involving $\ln(-k_*\tau)$ and $\ln^2(-k_*\tau)$ in (??) and (??)–(??). The analysis in this section shows that this agreement is not restricted to the lowest powers of $\ln(-k_*\tau)$ but will extend to the leading-logarithm terms at all orders.

Comparing the analysis of Elliston et al. [?] with the discussion given in this section highlights the advantages and disadvantages of the renormalization-group technique. Eqs. (??)–(??) involve combinations of the Riemann curvature tensor and its derivative which one could not expect to obtain by accident. The analysis of Elliston et al. [?] provides an explanation for the precise combinations which appear, and shows how they can be understood in geometrical terms. On its own, however, it does not demonstrate agreement to higher

orders in $\ln(-k_*\tau)$. This is important, because as we argued in §?? and was demonstrated explicitly in Fig. ??, contributions from all powers of $\ln(-k_*\tau)$ are required to obtain an accurate prediction.

In contrast, by combining the dynamical renormalization-group framework with a factorization theorem, we derived a correct set of transport equations directly from the linear $\ln(-k_*\tau)$ ‘divergences’. (As we explained in footnote ?? on p. ??, there is no need to take special measures to account for the double-logarithmic term.) Therefore the DRGE procedure gives an algorithmic method to obtain a correct evolution equation, valid to leading-logarithm order. But conversely it gives no information about the physical interpretation of this evolution, in the same way that the successful calculation of a negative β -function in QCD did not directly give an interpretation of asymptotic freedom as a consequence of antiscreening due to gluonic fluctuations.

In this case Elliston et al. were able to guess the correct physical interpretation of Eqs. (??) and (??)–(??). But it is easy to imagine that, in more complex circumstances (perhaps where the effect of time-dependence from loop corrections is included), it would no longer be possible to guess the correct evolution equation using simple physical arguments. In such cases the dynamical renormalization group provides a straightforward and algorithmic approach. In canonical models it requires the extra complexity of a next-order calculation to obtain the first contributions to the linear term in the Taylor expansion [that is, the term proportional to $\ln(-k_*\tau)$]. For the nontrivial metric, the relevant new terms involving the Riemann curvature appear already at lowest order.

6 Summary and discussion

The separate universe picture is a powerful method to compute the superhorizon evolution of inflationary correlation functions. Our main result is a derivation of this method from the renormalization group technique. We have focused on the bispectrum, where the simplest discussion can be given, but it is clear what would be involved in extending the analysis to higher n -point functions.

The main steps in the argument are simple. We first show that the vertex integral obeys a ‘factorization’ principle: to all orders in the slow-roll expansion, time-dependent logarithms appear only in combination with a few fixed functions of the external momenta. The second step is to derive renormalization-group equations for the coefficients of these fixed functions. These equations turn out to be the transport equations for the two- and three-point functions, which were shown in Ref. [?] to be equivalent to the separate-universe Taylor expansion introduced by Lyth & Rodríguez. Therefore, in combination, the factorization principle and RGE approach reproduce the separate universe method. Technically our argument demonstrates this in the leading-logarithm approximation.

Our result suggests a systematic method to incorporate more than the lowest-order contribution to each correlation function. Burgess et al. [?] studied time-dependent effects from loop-level corrections to the power spectrum, making use of the same dynamical renormalization group methods employed in §??. More recently, Collins et al. [?] studied effective field theories in a nonequilibrium or time-dependent setting, finding nontrivial contributions to the time-dependence after integrating out a massive field. To perform calculations analogous to these for n -point functions with $n \geq 3$ during an inflationary phase it would be necessary to

proceed as in §??—first, proving a suitable factorization theorem; and second, constructing the relevant renormalization-group equation. (For the two-point function, factorization is usually a much more straightforward question due to the absence of vertex integrals.)

The factorization property admits a clear physical interpretation which is related to the usual division of separate-universe calculations into a quantum contribution, generated near horizon-crossing, followed by classical evolution. The renormalization-group analysis gives a sharper version of this division. Fluctuations are created as they inflate beyond the Hubble length, on a timescale which is rapid compared to their subsequent evolution and which is insensitive to any model-dependent parameters. Factorization divides the physics of the fluctuations into two parts: a ‘hard’ model-independent subprocess associated with the creation event, and ‘soft’, model-dependent interactions which subsequently dress the hard subprocess. One way to understand why correlation functions factorize into self-contained parts associated with these ‘hard’ and ‘soft’ processes is that the disparity in timescales prevents interference effects between the two.

Because we work to tree-level, it follows that the soft physics can be described using the classical equations of motion. Had we chosen to compute beyond tree-level, the applicability of classical equations of motion might be lost but the division into hard and soft physics would remain. Therefore, as a conceptual tool, the hard/soft division takes the place of the usual quantum/classical separation, allowing correlation functions to be calculated without invoking any classical phase at all. Naturally, this does not absolve us of the obligation to explain why the correlation functions obtained in this way can subsequently be re-interpreted as the correlations of a classical stochastic process—that is, why we see fluctuations which behave classically. The renormalization-group method has nothing new to say about this difficulty.

In the course of our analysis we treated all fields as perturbatively massless. (That is, mass terms are treated as interactions.) This has two important consequences. First, all k -dependence occurs as integer power laws modified by logarithms. For example, $k^{3+\delta} \approx k^3[1 + \delta \ln k + O(\delta^2)]$; it is the infinite sum of logarithms which reproduces a power-law scaling. This restriction to integer powers of k was helpful in proving the factorization theorem in §??, which depended on the existence of a ‘gap’ in the asymptotic expansion of each wavefunction in integer powers of k , Eq. (??). In our applications, the coefficient δ is of order a slow-roll parameter, so in a strict mathematical interpretation we are restricted to scenarios in which the slow-roll parameters are small throughout the evolution. Second, the appearance of time-dependent logarithmic ‘divergences’ of the form $\ln(-k_*\tau)$ is itself a consequence of the massless approximation. Incorporation of masses would provide a cutoff which prevents these terms growing unboundedly.

Finally, it should be remembered that our analysis does not apply to all theories. Scenarios which contain other distinguished scales need not satisfy the conditions of the factorization theorem proved in §??. To obtain this theorem it was assumed that asymptotic information about the wavefunctions provided all necessary data about the superhorizon regime. Where this assumption fails, other combinations of the external momenta could receive significant enhancements. This corresponds to the well-known restriction that, in applying separate-universe arguments, all relevant scales should first be outside the horizon.

Acknowledgments

We would like to thank Enrico Pajer for discussions, and Joseph Elliston, Richard Holman and Andrew J. Tolley for comments on a draft version of this paper. We would like to thank the referee for drawing our attention to Ref. [?].

MD is financially supported by Fundação para a Ciência e a Tecnologia [grant SFRH/BD/44958/2008]. RHR acknowledges financial support from Fundação para a Ciência e a Tecnologia [grant SFRH/BD/35984/2007] in the early stages of this work. RHR also acknowledges the hospitality of the University of Sussex. DS acknowledges support from the Science and Technology Facilities Council [grant number ST/I000976/1] and the Leverhulme Trust. DS acknowledges that this material is based upon work supported in part by the National Science Foundation under Grant No. 1066293 and the hospitality of the Aspen Center for Physics. The research leading to these results has received funding from the European Research Council under the European Union’s Seventh Framework Programme (FP/2007–2013) / ERC Grant Agreement No. [308082].

DS would like to thank Misao Sasaki for the suggestion (made long ago at the workshop ‘Non-Gaussianity from Inflation’ held at DAMTP, Cambridge, in April 2006) that the methods of Refs. [? ?] could be used to determine the superhorizon evolution of the three-point function.

A Three-point function at next-to-leading order

In this Appendix we collect the components of the three-point function up to NLO in the slow-roll expansion, and briefly sketch some details of its calculation. For full details of the ‘in-in’ or ‘Schwinger–Keldysh’ method used to obtain inflationary correlation functions we refer to the literature [? ? ? ? ?].

A.1 Propagator

We require the second- and third-order action for a collection of light, canonically-normalized scalar fields coupled to gravity. The second-order action has already been given in (??). To obtain the propagator G , defined in conformal time by

$$G^{\alpha\beta}(\mathbf{x}, \tau; \tilde{\mathbf{x}}, \tilde{\tau}) = \langle \delta\phi^\alpha(\mathbf{x}, \tau) \delta\phi^\beta(\tilde{\mathbf{x}}, \tilde{\tau}) \rangle, \quad (\text{A.1})$$

we should invert the quadratic structure appearing in the action. That gives

$$\left\{ \delta_{\alpha\beta} \left(\frac{\partial^2}{\partial \tau^2} + 2 \frac{\partial \ln a}{\partial \tau} \frac{\partial}{\partial \tau} - \partial^2 \right) + a^2 m_{\alpha\beta} \right\} G^{\beta\gamma} = -\frac{i}{a^2} \delta_\alpha^\gamma \delta(\mathbf{x} - \tilde{\mathbf{x}}) \delta(\tau - \tilde{\tau}). \quad (\text{A.2})$$

Because spatial translations and rotations are unbroken by the de Sitter background, G can depend only on the relative displacement $\mathbf{x} - \tilde{\mathbf{x}}$. Diagonalizing this by passing to Fourier space,

$$G^{\alpha\beta}(\tau, \tilde{\tau}; \mathbf{x} - \tilde{\mathbf{x}}) = \int \frac{d^3 k}{(2\pi)^3} G_k^{\alpha\beta}(\tau, \tau') e^{i\mathbf{k} \cdot (\mathbf{x} - \tilde{\mathbf{x}})}, \quad (\text{A.3})$$

and introducing a dimensionless time $y = k\tau = -k/aH$, we find

$$\left\{ \delta_{\alpha\beta} \left(\frac{d^2}{dy^2} + 2\frac{a'}{a} \frac{d}{dy} + 1 \right) + \frac{a^2}{k^2} m_{\alpha\beta} \right\} G_k^{\beta\gamma} = -\frac{i}{ka^2} \delta_\alpha^\gamma \delta(y - \tilde{y}), \quad (\text{A.4})$$

where $\tilde{y} = k\tilde{\tau}$ and a' now denotes da/dx . Eq. (??) can be transformed into a Bessel equation by making the substitution $G_k^{\alpha\beta} = w^{\alpha\beta} \sqrt{-y}/a$. After expanding a to next-order (for details, see Refs. [? ?]) we finally arrive at an equation which can be solved perturbatively,

$$\left\{ \delta_{\alpha\beta} \left(\frac{d^2}{dy^2} + \frac{1}{y} \frac{d}{dy} + \left[1 - \frac{9/4}{y^2} \right] \right) - \frac{3M_{\alpha\beta}}{y^2} \right\} w^{\beta\gamma} = -\frac{i}{ka\sqrt{-y}} \delta_\alpha^\gamma \delta(y - \tilde{y}), \quad (\text{A.5})$$

where $M_{\alpha\beta} = \varepsilon_* \delta_{\alpha\beta} + u_{\alpha\beta}$ was defined in Eq. (??). It plays the role of an effective mass term, including contributions from coupling to the background geometry.

Homogeneous solution. The technology to solve Eq. (??) perturbatively in $M_{\alpha\beta}$ was developed by Nakamura & Stewart [?] and Gong & Stewart [? ?]. We adopt their methods, with some small modifications noted below to keep each wavefunction purely imaginary in the late-time limit. The first step is to find a solution $v^{\beta\gamma}$ of the homogeneous equation,

$$\left\{ \delta_{\alpha\beta} \left(\frac{d^2}{dy^2} + \frac{1}{y} \frac{d}{dy} + \left[1 - \frac{9/4}{y^2} \right] \right) - \frac{3M_{\alpha\beta}}{y^2} \right\} v^{\beta\gamma} = 0. \quad (\text{A.6})$$

To do so we work perturbatively in $M_{\alpha\beta}$, writing $v^{\alpha\beta} = v_0^{\alpha\beta} + v_1^{\alpha\beta} + \dots$, with $v_n^{\alpha\beta} = \mathcal{O}(M^n)$. Up to $\mathcal{O}(M)$ we find

$$\delta_{\alpha\beta} \left(\frac{d^2}{dy^2} + \frac{1}{y} \frac{d}{dy} + \left[1 - \frac{9/4}{y^2} \right] \right) v_0^{\beta\gamma} = 0, \quad (\text{A.7a})$$

$$\delta_{\alpha\beta} \left(\frac{d^2}{dy^2} + \frac{1}{y} \frac{d}{dy} + \left[1 - \frac{9/4}{y^2} \right] \right) v_1^{\beta\gamma} = \frac{3}{y^2} M_{\alpha\beta} v_0^{\beta\gamma}. \quad (\text{A.7b})$$

Eq. (??) has solution in terms of Bessel functions of order $3/2$. The linear combination with correct boundary conditions to match the Minkowski-space wavefunctions in the subhorizon limit $k/aH = |y| \rightarrow \infty$ is the Hankel function,

$$v_0^{\alpha\beta} = \mathcal{U}^{\alpha\beta} H_{3/2}^{(2)}(-k\tau) \quad (\text{A.8})$$

where $\mathcal{U}^{\alpha\beta}$ is an arbitrary constant matrix. Eq. (??) can be solved using the causal Green's function $\Gamma(y, z)$ constructed by Gong & Stewart [? ?]. This satisfies

$$\frac{d^2\Gamma}{dy^2} + \frac{1}{y} \frac{d\Gamma}{dy} + \left[1 - \frac{9/4}{y^2} \right] \Gamma = \delta(y - z), \quad (\text{A.9})$$

and can be written explicitly in terms of the Hankel functions

$$\Gamma(y, z) = -\frac{i\pi z}{4} \times \begin{cases} 0 & y < z \\ H_{3/2}^{(2)}(-y)H_{3/2}^{(1)}(-y) - H_{3/2}^{(1)}(-y)H_{3/2}^{(2)}(-y) & y > z \end{cases}. \quad (\text{A.10})$$

After integrating over Γ , the solution $v_1^{\alpha\beta}$ can be written

$$v_1^{\alpha\beta} = M^\alpha_\gamma \mathcal{U}^{\gamma\beta} v_1, \quad (\text{A.11})$$

where v_1 satisfies [? ?]

$$v_1 = i\sqrt{\frac{2}{\pi}} \frac{1}{(-y)^{3/2}} \left(2e^{iy} - (1+iy)e^{-ix} \int_{-\infty}^y \frac{dz}{z} e^{2iz} \right). \quad (\text{A.12})$$

In writing (??) we have adjusted the position of the indices on M to match on both sides of the equation. However, since we are generally working with a flat field-space metric, except in §??, there is no distinction between co- and contravariant indices.

The scalar part of the zero-order wavefunction $H_{3/2}^{(2)}(-y)$ has the property that it is purely imaginary at late times,

$$H_{3/2}^{(2)}(-y) \rightarrow i\sqrt{\frac{2}{\pi}} \frac{1}{(-y)^{3/2}} + \mathcal{O}(y). \quad (\text{A.13})$$

However, in this limit v_1 is complex. Assembling (??), (??) and (??) gives

$$v^{\alpha\beta} \rightarrow i\sqrt{\frac{2}{\pi}} \frac{1}{(-y)^{3/2}} \left(\mathcal{U}^{\alpha\beta} + M^{\alpha}_{\gamma} \mathcal{U}^{\gamma\beta} \left[2 - \gamma_E - \ln(-2y) - i\frac{\pi}{2} \right] \right). \quad (\text{A.14})$$

It is the term involving $i\pi/2$ which is problematic. This arises from the asymptotic expansion of the exponential integral appearing in Eq. (??). We will see below that to clearly exhibit the cancellation of unwanted divergent terms in the calculation of the three-point functions it is helpful to adjust the phase of $v^{\alpha\beta}$ so that this term is removed. Therefore we redefine $v_{\alpha\beta}$ by a pure phase

$$v^{\alpha\beta} \rightarrow \tilde{v}^{\alpha\beta} = \exp(i\theta^{\alpha\gamma}) v^{\gamma\beta}, \quad (\text{A.15})$$

where $\theta^{\alpha\beta}$ is a constant real matrix. This redefinition does *not* change the propagator, which depends on the combination $\mathbf{v}^\dagger \cdot \mathbf{v}$ in which this phase cancels out. To remove the imaginary term in Eq. (??) we should choose $\theta^{\alpha\beta} = (\pi/2)M^{\alpha\beta}$. The net effect is to replace v_1 in Eq. (??) with a redefined function (cf. Refs. [? ? ?])

$$v_1 \rightarrow \tilde{v}_1 = i\sqrt{\frac{2}{\pi}} \frac{1}{(-y)^{3/2}} \left(2e^{iy} + i\frac{\pi}{2}(1-iy)e^{iy} - (1+iy)e^{-iy} \int_{-\infty}^y \frac{dz}{z} e^{2iz} \right). \quad (\text{A.16})$$

Feynman propagator. We now return to Eq. (??) and use the homogeneous solution $v^{\alpha\beta}$ to construct the Feynman Green's function $w^{\alpha\beta}$.²⁰ In this section it is helpful to switch to matrix notation, with $(\mathbf{v})_{\alpha\beta} = v^{\alpha\beta}$, $(\mathbf{w})_{\alpha\beta} = w^{\alpha\beta}$ and so on. In the absence of the mass matrix \mathbf{M} the field fluctuations are uncorrelated, which implies that we must choose $\mathbf{U} = \mathbf{1}$. Together, the Feynman boundary conditions and hermiticity require

$$\mathbf{u} = \begin{cases} \mathbf{v}^\dagger(\tilde{y}) \cdot \mathbf{Y} \cdot \mathbf{v}(y) & y < \tilde{y} \\ \mathbf{v}^\dagger(y) \cdot \mathbf{Y} \cdot \mathbf{v}(\tilde{y}) & y > \tilde{y} \end{cases} \quad (\text{A.17})$$

where \mathbf{Y} is an Hermitian matrix. We will take it to commute with \mathbf{v} and \mathbf{v}^\dagger . This can be justified (with our choice $\mathbf{U} = \mathbf{1}$) because Eq. (??) implies that both \mathbf{v} and \mathbf{v}^\dagger are power series

²⁰Strictly, in the context of the Schwinger method this is the time-ordered $++$ correlation function. Its complex conjugate is the time-ordered $--$ correlation function. The $+-$ and $-+$ correlation functions can be extracted from the two alternatives in Eq. (??), bearing in mind that $-$ fields are later than $+$ fields for the purpose of time ordering, no matter at what physical time they are evaluated.

in the Hermitian matrix \mathbf{M} . The normalization condition in (??) likewise guarantees that \mathbf{Y} is a power series in \mathbf{M} . Therefore all these matrices commute. Imposing this normalization condition, we conclude

$$\mathbf{Y} \cdot \left(\mathbf{v}(\tilde{y}) \cdot \frac{d\mathbf{v}^\dagger(\tilde{y})}{d\tilde{y}} - \mathbf{v}^\dagger(\tilde{y}) \cdot \frac{d\mathbf{v}(\tilde{y})}{d\tilde{y}} \right) = -\frac{i}{(-\tilde{y})^{1/2}ka(\tilde{y})} \times \mathbf{1}. \quad (\text{A.18})$$

The matrix in brackets (\dots) is the ‘Wronskian matrix’ for \mathbf{v} and \mathbf{v}^\dagger . Abel’s theorem shows that it can be written

$$\mathbf{W}(\tilde{y}) = \mathbf{v}(\tilde{y}) \cdot \frac{d\mathbf{v}^\dagger(\tilde{y})}{d\tilde{y}} - \mathbf{v}^\dagger(\tilde{y}) \cdot \frac{d\mathbf{v}(\tilde{y})}{d\tilde{y}} = \mathbf{W}(\tilde{y}_0) \exp \left(-\int_{\tilde{y}_0}^{\tilde{y}} \frac{dz}{z} \right) = \mathbf{W}(\tilde{y}_0) \frac{\tilde{y}_0}{\tilde{y}}, \quad (\text{A.19})$$

where \tilde{y}_0 is arbitrary and $\mathbf{W}(\tilde{y}_0)$ is independent of \mathbf{M} . With our choice $\mathcal{U} = 1$, we find

$$\mathbf{W}(\tilde{y}) = -\frac{4i}{\pi(-\tilde{y})} \times \mathbf{1}. \quad (\text{A.20})$$

Therefore \mathbf{Y} is Hermitian and commutes with \mathbf{M} , in agreement with our earlier argument. In conclusion, the Feynman propagator can be written

$$\mathbf{G}_k = \frac{\pi}{4k} \frac{(-y)^{1/2}(-\tilde{y})^{1/2}}{a(y)a(\tilde{y})} \times \begin{cases} \mathbf{v}^\dagger(\tilde{y}) \cdot \mathbf{v}(y) & y < \tilde{y} \\ \mathbf{v}^\dagger(y) \cdot \mathbf{v}(\tilde{y}) & y > \tilde{y} \end{cases}. \quad (\text{A.21})$$

Uniform expansion of \mathbf{G}_k to next-order. We require the lowest-order and next-order contributions in slow-roll to \mathbf{G}_k . This can be obtained from the solutions $v_0^{\alpha\beta}$ and $v_1^{\alpha\beta}$ together with next-order corrections to the scale factor a , for which purpose we must Taylor expand around an arbitrary reference scale k_* as described in §??. Denoting evaluation at the horizon-crossing time for k_* by a subscript ‘*’, we find that \mathbf{G}_k behaves like the product of two effective wave-function factors \mathbf{w} ,

$$\mathbf{G}_k = \begin{cases} \mathbf{w}^\dagger(\tilde{y}) \cdot \mathbf{w}(y) & y < \tilde{y} \\ \mathbf{w}^\dagger(y) \cdot \mathbf{w}(\tilde{y}) & y > \tilde{y} \end{cases}, \quad (\text{A.22})$$

where $\mathbf{w} = \mathbf{w}_0 + \mathbf{w}_1 + \dots$ satisfies

$$\mathbf{w}_0(y) = \frac{i}{\sqrt{2k^3}} H_*(1 - iy) e^{iy} \cdot \mathbf{1} \quad (\text{A.23a})$$

$$\begin{aligned} \mathbf{w}_1(y) = \frac{i}{\sqrt{2k^3}} H_* \Big\{ \varepsilon_* \left[\left(\ln \frac{k_*}{k} (-y) \right) - 1 \right] (1 - iy) e^{iy} \cdot \mathbf{1} \\ + \mathbf{M} \cdot \left[2e^{iy} + i \frac{\pi}{2} (1 - iy) e^{iy} - (1 + iy) e^{-iy} \int_{-\infty}^y \frac{dz}{z} e^{2iz} \right] \Big\}. \end{aligned} \quad (\text{A.23b})$$

[This result was used in the main text in Eq. (??).] Eqs. (??) and (??) are both exact in y but perturbative in powers of slow-roll and mass M .

A.2 Asymptotic behaviour of wavefunctions

In practical calculations we require the late-time expansion of the elementary wavefunctions, \mathbf{w} , in order to compute the contribution from external lines. In the present case it is only necessary to compute these up to $\mathcal{O}(y^2)$. We find

$$\mathbf{w}(y) = \frac{i}{\sqrt{2k^3}} H_* e^{iy} \left\{ \mathbf{1} + \mathbf{A}_* - \mathbf{u} \ln \frac{k_*}{k} (-y) - iy(\mathbf{1} + \mathbf{A}_*) + iy\mathbf{u} \ln \frac{k_*}{k} (-y) + \mathcal{O}(y)^2 \right\}, \quad (\text{A.24})$$

where

$$\mathbf{A} \equiv \varepsilon_* \left(1 - \gamma_E - \ln \frac{2k}{k_*} \right) \mathbf{1} + \mathbf{u}_* \left(2 - \gamma_E - \ln \frac{2k}{k_*} \right), \quad (\text{A.25})$$

and \mathbf{u} is defined in Eq. (??).

A.3 Three-point function

To compute the three-point function we require the action to third-order. This calculation is well-documented in the literature; see, for examples, Refs. [? ?]. Fluctuations in the fields will mix with the metric, which can be accommodated using the ADM decomposition (??)

$$ds^2 = -N^2 dt^2 + h_{ij}(dx^i + N^i dt)(dx^j + N^j dt), \quad (\text{A.26})$$

where the lapse N and shift vector N^i are determined by constraint equations which follow from invariance under temporal and spatial reparametrizations. At the level of the background we have $N = 1$ and $N^i = 0$. When perturbations are included there are order-by-order corrections, discussed in §??,

$$N = 1 + \sum_{n=1}^{\infty} \alpha_n, \quad (\text{A.27a})$$

$$N_i = \sum_{n=1}^{\infty} \partial_i \vartheta_n + \text{solenoidal component}. \quad (\text{A.27b})$$

The solenoidal part of N_i is important only when computing the action at fourth-order and above [? ?]. (Spatial indices i, j, \dots are raised and lowered using the three-metric h_{ij} .) Also, although in principle there could be contributions to the third-order action from α_n and ϑ_n with $n \leq 3$, the structure of the constraints in the action guarantees that only α_1 and ϑ_1 contribute [? ?]. Specializing to the spatially flat gauge $h_{ij} = a^2 \delta_{ij}$, we find²¹

$$\begin{aligned} S_{\phi\phi\phi} \supseteq \frac{1}{2} \int d^3x dt a^3 \Bigg\{ & -\frac{1}{3} V_{\alpha\beta\gamma} \delta\phi^\alpha \delta\phi^\beta \delta\phi^\gamma - \frac{2}{a^2} \delta\dot{\phi}^\alpha \partial_j \vartheta_1 \partial_j \delta\phi_\alpha \\ & + \alpha_1 \left[-\frac{1}{a^2} \partial_j \delta\phi^\alpha \partial_j \delta\phi_\alpha - \delta\dot{\phi}^\alpha \delta\dot{\phi}_\alpha - V_{\alpha\beta} \delta\phi^\alpha \delta\phi^\beta \right. \\ & \quad \left. - \frac{1}{a^4} \partial_i \partial_j \vartheta_1 \partial_i \partial_j \vartheta_1 + \frac{1}{a^4} \partial^2 \vartheta_1 \partial^2 \vartheta_1 + \frac{2}{a^2} \dot{\phi}^\alpha \partial_j \vartheta_1 \partial_j \delta\phi_\alpha \right. \\ & \quad \left. + \alpha_1 \left(2\dot{\phi}^\alpha \delta\dot{\phi}_\alpha + \frac{4H}{a^2} \partial^2 \vartheta_1 + \alpha_1 (6H^2 - \dot{\phi}^2) \right) \right] \Bigg\} \end{aligned} \quad (\text{A.28})$$

²¹Eq. (??) contains a term linear in $\partial^2 \vartheta_1 / a^2$, in apparent contradiction with the conclusions of §??. This discrepancy is only an appearance, however. Eq. (??) has been simplified using the constraints, which allow us to trade ϑ_1 for specific combinations of the other metric and scalar fluctuations. Therefore any appearance of nonlocality in (??) is fictitious and will disappear in the final answer. See the discussion on p. ??.

where $\dot{\phi}^2 = \dot{\phi}_\alpha \dot{\phi}_\alpha$ and repeated indices in the lower position are summed using the flat Euclidean metric δ_{ij} . Eq. (??) is exact. It has been obtained using an expansion in the amplitudes $|\delta\phi|$ but does not invoke an expansion in slow-roll parameters. The solution for α_1 is

$$\alpha_1 = \frac{\dot{\phi}_\alpha \delta\phi_\alpha}{2H}, \quad (\text{A.29})$$

and the solution for ϑ_1 was used in (??),

$$\frac{4H}{a^2} \partial^2 \vartheta_1 = -2V_\alpha \delta\phi_\alpha - 2\dot{\phi}_\alpha \delta\dot{\phi}_\alpha + 2\alpha_1(-6H^2 + \dot{\phi}^2). \quad (\text{A.30})$$

We wish to compute the three-point function up to terms which are next-to-leading (‘NLO’) in a fixed-order slow-roll expansion. For this purpose we need only extract the LO and NLO terms in (??). The LO contributions were given in Ref. [?] , where the corresponding three-point function was also obtained. They are

$$S_{\phi\phi\phi}^{\text{LO}} \supseteq \frac{1}{2} \int d^3x dt a^3 \left(\frac{1}{H} \dot{\phi}^\beta \delta\dot{\phi}^\alpha \partial_j \delta\phi_\alpha \partial_j \partial^{-2} \delta\dot{\phi}_\beta - \alpha_1 \left[\delta\dot{\phi}^\alpha \delta\dot{\phi}_\alpha + \frac{1}{a^2} \partial_j \delta\phi^\alpha \partial_j \delta\phi_\alpha \right] \right), \quad (\text{A.31})$$

(In Ref. [?] the action was quoted in a different but equivalent form.) The new contributions at NLO are

$$\begin{aligned} S_{\phi\phi\phi}^{\text{NLO}} \supseteq \frac{1}{2} \int d^3x dt a^3 & \left(g_{\alpha\beta\gamma} H^2 \delta\phi^\alpha \delta\phi^\beta \delta\phi^\gamma - \frac{1}{2} q^\alpha H \delta\dot{\phi}^\beta \partial_j \delta\phi_\beta \partial_j \partial^{-2} \delta\phi_\alpha \right. \\ & - \frac{\dot{\phi}^\alpha \dot{\phi}^\beta \dot{\phi}^\gamma}{2H^2} \delta\phi_\alpha \partial_j \delta\phi_\beta \partial_j \partial^{-2} \delta\dot{\phi}_\gamma \\ & \left. - \frac{\dot{\phi}^\alpha \dot{\phi}^\beta \dot{\phi}^\gamma}{8H^3} \delta\phi_\alpha \left[\partial_i \partial_j \partial^{-2} \delta\dot{\phi}_\beta \partial_i \partial_j \partial^{-2} \delta\dot{\phi}_\gamma - \delta\dot{\phi}_\beta \dot{\phi}_\gamma \right] \right), \end{aligned} \quad (\text{A.32})$$

where the coefficient tensors $g_{\alpha\beta\gamma}$ and q_α satisfy

$$g_{\alpha\beta\gamma} = -\frac{1}{3} \frac{V_{\alpha\beta\gamma}}{H^2} - \frac{\dot{\phi}_\alpha}{2H} \frac{V_{\beta\gamma}}{H^2} + \frac{3}{4} \frac{\dot{\phi}_\alpha \dot{\phi}_\beta \dot{\phi}_\gamma}{H^3} \quad (\text{A.33a})$$

$$q^\alpha = 2 \frac{\ddot{\phi}^\alpha}{H^2} + 2\varepsilon \frac{\dot{\phi}^\alpha}{H} = 2 \frac{\dot{\phi}_\beta}{H} u^{\alpha\beta}. \quad (\text{A.33b})$$

(We are using $\varepsilon = \dot{\phi}^2/2H^2$, which is accurate at lowest order in the slow-roll approximation.) In particular, $g_{\alpha\beta\gamma}$ is symmetric under the exchange $\beta \leftrightarrow \gamma$, but not under exchanges involving α . The three-point function for the $V_{\alpha\beta\gamma}$ term was obtained by Falk, Rangarajan & Srednicki [?] , and later recomputed by Zaldarriaga [?] . These calculations enable the three-point function for $g_{\alpha\beta\gamma}$ to be deduced,²² but results for the remaining terms in Eq. (??) have not yet been obtained.

A.3.1 Next-to-leading interactions

Terms in the NLO action (??) contribute to the three-point function beginning at next-to-leading order in the slow-roll expansion. Therefore, we need only retain the lowest-order

²²The three-point functions obtained by Falk et al. and Zaldarriaga contained a small typographical error, which was corrected in Ref. [?] . The correct expression is given in Eq. (??).

contribution (??) to the propagator. Also we require only the lowest order estimate $a = -(H_*\tau)^{-1}$ for the scale factor, where as usual we have introduced an arbitrary reference scale k_* which serves as the basis for a Taylor expansion of all background quantities, as explained in §??.

Notation. The full set of three-point functions is lengthy. To simplify the presentation, it is helpful to factor out a number of common elements. In particular, we define each three-point function in terms of a quantity $\mathcal{B}^{\alpha\beta\gamma}$, in which we extract a further common factor of H_*^4 in comparison to $B^{\alpha\beta\gamma}$ defined in (??). This must be symmetrized to produce the final result,

$$\langle \delta\phi_\alpha(\mathbf{k}_1)\delta\phi_\beta(\mathbf{k}_2)\delta\phi_\gamma(\mathbf{k}_3) \rangle_\tau = (2\pi)^3 \delta(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3) \frac{H_*^4}{4 \prod_i k_i^3} \mathcal{B}_{\alpha\beta\gamma}(k_1, k_2, k_3) \quad (\text{A.34})$$

$$+ [5 \text{ permutations}],$$

where the permutations are formed by simultaneously exchanging the index and momentum pairs (α, \mathbf{k}_1) , (β, \mathbf{k}_2) and (γ, \mathbf{k}_3) . We also define

$$k_t = k_1 + k_2 + k_3, \quad (\text{A.35a})$$

$$K^2 = \sum_{i < j} k_i k_j = k_1 k_2 + k_1 k_3 + k_2 k_3, \quad (\text{A.35b})$$

and set γ_E to be the Euler–Mascheroni constant, $\gamma_E \approx 0.577$.

We highlight terms involving $\ln(-k_*\tau)$ in **green**. For vertex and internal wavefunction corrections, these contribute to the divergent part of the three-point function (??). For external wavefunctions, they contribute to (??).

$\delta\phi^\alpha\delta\phi^\beta\delta\phi^\gamma$ **operator**

$$\mathcal{B}_{\alpha\beta\gamma} \supseteq \frac{g_{\alpha\beta\gamma}^*}{2} \left(\frac{4}{9} k_t^2 - K^2 k_t - \frac{1}{3} \left(\sum_i k_i^3 \right) \left[\ln(-k_*\tau) + \ln \frac{k_t}{k_*} + \gamma_E + \frac{1}{3} \right] \right) \quad (\text{A.36})$$

$\delta\dot{\phi}^\beta\partial_j\delta\phi_\beta\partial_j\partial^{-2}\delta\phi_\alpha$ **operator**

$$\mathcal{B}_{\alpha\beta\gamma} \supseteq \frac{q_\gamma^* \delta_{\alpha\beta}}{4} \frac{k_1^2}{k_3^2} (\mathbf{k}_2 \cdot \mathbf{k}_3) \left(k_t - \frac{k_2 k_3}{k_t} - k_1 \left[\ln(-k_*\tau) + \ln \frac{k_t}{k_*} + \gamma_E \right] \right) \quad (\text{A.37})$$

$\delta\phi_\alpha\partial_j\delta\phi_\beta\partial_j\partial^{-2}\delta\dot{\phi}_\gamma$ **operator**

$$\mathcal{B}_{\alpha\beta\gamma} \supseteq \frac{\dot{\phi}_\alpha^* \dot{\phi}_\beta^* \dot{\phi}_\gamma^*}{4H_*^3} (\mathbf{k}_2 \cdot \mathbf{k}_3) \left(k_t - \frac{k_1 k_2}{k_t} - k_3 \left[\ln(-k_*\tau) + \ln \frac{k_t}{k_*} + \gamma_E \right] \right) \quad (\text{A.38})$$

$\delta\phi_\alpha\delta\dot{\phi}_\beta\delta\dot{\phi}_\gamma$ **operators**

$$\mathcal{B}_{\alpha\beta\gamma} \supseteq -\frac{\dot{\phi}_\alpha^* \dot{\phi}_\beta^* \dot{\phi}_\gamma^*}{16H_*^3} \frac{1}{k_t} \left(1 + \frac{k_1}{k_t} \right) ([\mathbf{k}_2 \cdot \mathbf{k}_3]^2 - k_2^2 k_3^2) \quad (\text{A.39})$$

A.3.2 Leading-order interactions

The LO operators which appear in (??) contribute to the three-point function at *both* lowest-order and next-order. The next-order terms come from use of (??) to correct the wavefunctions associated with each internal or external line, or from subleading terms in the Taylor expansion of background quantities such as H and $\dot{\phi}^\alpha$, including factors of H which appear in the scale factor.

Notation. These sources of next-order corrections were identified and described by Chen et al. [?]. Subsequently, rules for explicit evaluation of the integrals which appear at intermediate stages of the computation were given by Burrage et al. [? ?], to which we refer for all technical details. For our present purposes we require only the functions

$$J_0(k_3) = \frac{1}{\vartheta_3} \ln(1 - \vartheta_3) \quad (\text{A.40a})$$

$$J_1(k_3) = \frac{1}{\vartheta_3^2} \left(\vartheta_3 + \ln[1 - \vartheta_3] \right), \quad (\text{A.40b})$$

where ϑ_3 is defined by

$$\vartheta_3 = 1 - \frac{2k_3}{k_t}. \quad (\text{A.40c})$$

In writing Eqs. (??)–(??) we have used the convention that, although each expression depends on all k_i , only the asymmetrically occurring momentum is written explicitly. As described in Refs. [? ?], the $J_n(k_3)$ approach finite values in the limit $\vartheta_3 \rightarrow 0$ which corresponds to either of the (physical) squeezed configurations $k_1 \rightarrow 0$ or $k_2 \rightarrow 0$.

To describe contributions from the late-time limit of (??), which are generated by external wavefunction factors, we use Eq. (??) and the combination

$$t_\alpha = 3\varepsilon \frac{\dot{\phi}_\alpha}{H} + \frac{\ddot{\phi}_\alpha}{H^2} = \frac{q_\alpha}{2} + 2\varepsilon \frac{\dot{\phi}_\alpha}{H}. \quad (\text{A.41})$$

As above, the following expressions should be symmetrized over simultaneous exchange of the index and momentum pairs (α, \mathbf{k}_1) , (β, \mathbf{k}_2) and (γ, \mathbf{k}_3) .

- $\delta\phi_\beta \delta\dot{\phi}^\alpha \delta\dot{\phi}_\alpha$ operator

LO plus NLO from external wavefunctions

$$\begin{aligned} \mathcal{B}_{\alpha\beta\gamma} \supseteq & -\frac{1}{4} \frac{k_2^2 k_3^2}{k_t} \frac{\dot{\phi}_*^\lambda}{H_*} \left(1 + \frac{k_1}{k_t} \right) \left(\delta_{\alpha\lambda} \delta_{\beta\gamma} + \left[A_{\alpha\lambda}^*(k_1) - \textcolor{teal}{u}_{\alpha\lambda}^* \ln(-k_* \tau) \right] \delta_{\beta\gamma} \right. \\ & \left. + \left[A_{\beta\gamma}^*(k_2) + A_{\beta\gamma}^*(k_3) - 2\textcolor{teal}{u}_{\beta\gamma}^* \ln(-k_* \tau) \right] \delta_{\alpha\lambda} \right) \end{aligned} \quad (\text{A.42})$$

NLO from vertex

$$\mathcal{B}_{\alpha\beta\gamma} \supseteq -\frac{1}{4} \frac{k_2^2 k_3^2}{k_t} \delta_{\beta\gamma} \left\{ 2\varepsilon_* \frac{\dot{\phi}_\alpha}{H_*} \left(1 + \frac{k_1}{k_t} \right) + t_\alpha^* \left[-\frac{k_1}{k_t} + \left(\gamma_E + \ln \frac{k_t}{k_*} \right) \left(1 + \frac{k_1}{k_t} \right) \right] \right\} \quad (\text{A.43})$$

NLO from internal wavefunctions

$$\begin{aligned} \mathcal{B}_{\alpha\beta\gamma} \supseteq & -\frac{1}{4} \frac{\dot{\phi}_\alpha^*}{H_*} \left\{ \textcolor{teal}{u}_{\beta\gamma}^* \left[(k_2^3 + k_3^3) \left(\ln(-k_* \tau) + \ln \frac{k_t}{k_*} + \gamma_E \right) - k_t (k_2^2 + k_3^2) \right] \right. \\ & - 3\varepsilon_* \delta_{\beta\gamma} \frac{k_2^2 k_3^2}{k_t} \left(\gamma_E + \ln \frac{k_t}{k_*} \right) \left(1 + \frac{k_1}{k_t} \right) \\ & \left. - M_{\beta\gamma}^* \frac{k_2^2 k_3^2}{k_t} \left[J_0(k_2) + J_0(k_3) + \frac{k_1}{k_t} \{ J_1(k_2) + J_1(k_3) \} \right] \right\} \\ & + \frac{1}{4} \frac{k_2^2 k_3^2}{k_t} \frac{\dot{\phi}_*^\lambda}{H_*} \left\{ \delta_{\beta\gamma} \left[M_{\alpha\lambda}^* \left(J_0(k_1) - \frac{k_1}{k_t} J_1(k_1) \right) - 2u_{\alpha\lambda}^* \right] + \delta_{\alpha\lambda} \left[2\varepsilon_* \delta_{\beta\gamma} - \frac{k_1(k_2 + k_3)}{k_2 k_3} u_{\beta\gamma}^* \right] \right\} \end{aligned} \quad (\text{A.44})$$

- $\delta\dot{\phi}^\alpha \partial_j \delta\phi_\alpha \partial_j \partial^{-2} \delta\dot{\phi}_\beta$ operator

LO plus NLO from external wavefunctions

$$\mathcal{B}_{\alpha\beta\gamma} \supseteq \frac{1}{2} k_2^2 \frac{\mathbf{k}_1 \cdot \mathbf{k}_3}{k_t} \frac{\dot{\phi}_*^\lambda}{H_*} \left(1 + \frac{k_3}{k_t} \right) \left(\delta_{\alpha\lambda} \delta_{\beta\gamma} + \left[A_{\alpha\lambda}^*(k_1) - u_{\alpha\lambda}^* \ln(-k_* \tau) \right] \delta_{\beta\gamma} \right. \\ \left. + \left[A_{\beta\gamma}^*(k_2) + A_{\beta\gamma}^*(k_3) - 2u_{\beta\gamma}^* \ln(-k_* \tau) \right] \delta_{\alpha\lambda} \right) \quad (\text{A.45})$$

NLO from vertex

$$\mathcal{B}_{\alpha\beta\gamma} \supseteq \frac{k_2^2 (\mathbf{k}_1 \cdot \mathbf{k}_3)}{2k_t} \delta_{\beta\gamma} \left\{ 2\varepsilon_* \frac{\dot{\phi}_*^\alpha}{H_*} \left(1 + \frac{k_3}{k_t} \right) - t_\alpha^* \left[\frac{k_3}{k_t} - \left(\gamma_E + \ln \frac{k_t}{k_*} \right) \left(1 + \frac{k_3}{k_t} \right) \right] \right\} \quad (\text{A.46})$$

NLO from internal wavefunctions

$$\mathcal{B}_{\alpha\beta\gamma} \supseteq \frac{\mathbf{k}_1 \cdot \mathbf{k}_3}{k_1^2} \frac{\dot{\phi}_*^\lambda}{2H_*} \left(\frac{k_1^2 k_2^2}{k_t} \left[-\delta_{\alpha\lambda} M_{\beta\gamma}^* [J_0(k_2) + J_0(k_3) + \frac{k_3}{k_t} \{J_1(k_2) - J_1(k_3)\}] \right. \right. \\ \left. \left. - \delta_{\beta\gamma} M_{\alpha\lambda}^* [J_0(k_1) + \frac{k_3}{k_t} J_1(k_1)] \right] - k_t \{k_1^2 \delta_{\gamma\lambda} u_{\alpha\beta}^* + k_2^2 \delta_{\beta\gamma} u_{\alpha\lambda}^*\} \right. \\ \left. - \left[\gamma_E + \ln \left(\frac{k_t}{k_*} \right) + \ln(-k_* \tau) \right] C_{\alpha\beta\gamma\lambda}^* \right. \\ \left. + \frac{1}{k_t} \left(\gamma_{\alpha\beta\gamma\lambda}^{0*} - \frac{1}{k_t} \{ \gamma_{\alpha\beta\gamma\lambda}^{1*} + \delta_{\alpha\beta\gamma\lambda}^{1*} \} \right) - \frac{1}{k_t} \left(\gamma_E + \ln \frac{k_t}{k_*} \right) \delta_{\alpha\beta\gamma\lambda}^{0*} \left(1 + \frac{k_3}{k_t} \right) \right\} \quad (\text{A.47})$$

where we have defined

$$\begin{cases} C_{\alpha\beta\gamma\lambda}^* = -k_1^3 u_{\beta\gamma}^* \delta_{\alpha\lambda} - k_2^3 \delta_{\beta\gamma} u_{\alpha\lambda}^* \\ \gamma_{\alpha\beta\gamma\lambda}^{0*} = -2k_1^2 k_3^2 \varepsilon_* \delta_{\beta\gamma} \delta_{\alpha\lambda} + k_1^2 k_2 k_3 \delta_{\alpha\lambda} u_{\beta\gamma}^* + k_2^2 k_3^2 \delta_{\alpha\lambda} (\varepsilon_* \delta_{\beta\gamma} + 2u_{\beta\gamma}^*) + k_1 k_2^2 k_3 \delta_{\beta\gamma} u_{\alpha\lambda}^* \\ \gamma_{\alpha\beta\gamma\lambda}^{1*} = k_2^2 k_3 (k_3^2 + k_1^2) \varepsilon_* \delta_{\alpha\lambda} \delta_{\beta\gamma} \\ \delta_{\alpha\beta\gamma\lambda}^{0*} = 3k_1^2 k_2^2 \varepsilon_* \delta_{\alpha\lambda} \delta_{\beta\gamma} \end{cases} \quad (\text{A.48})$$

- $\delta\phi_\beta(\partial_i \delta\phi_\alpha)^2$ operator.

This operator is simplified by our choice of phase convention (??). With the unmodified wavefunction (??), fast power-law divergences are generated by the internal wavefunctions which cancel against other power-law divergences arising from their external counterparts.

With the redefined phase in (??), all power-law divergences cancel locally, in the sense that we do not need to group terms from different parts of the calculation. In practice this is a considerable convenience.

LO plus NLO from external wavefunctions

$$\mathcal{B}_{\alpha\beta\gamma} \supseteq \frac{\mathbf{k}_2 \cdot \mathbf{k}_3}{4} \left(-\frac{\dot{\phi}_*^\lambda}{H_*} \frac{1}{k_t} \left(-k_t^2 + K^2 + \frac{k_1 k_2 k_3}{k_t} \right) \left[\delta_{\alpha\lambda} \delta_{\beta\gamma} + \{ A_{\alpha\lambda}^*(k_1) - u_{\alpha\lambda}^* \ln(-k_* \tau) \} \delta_{\beta\gamma} \right. \right. \\ \left. \left. + \{ A_{\beta\gamma}^*(k_2) + A_{\beta\gamma}^*(k_3) - 2u_{\beta\gamma}^* \ln(-k_* \tau) \} \delta_{\alpha\lambda} \right] \right) \quad (\text{A.49})$$

NLO from vertex

$$\mathcal{B}_{\alpha\beta\gamma} \supseteq \frac{(\mathbf{k}_2 \cdot \mathbf{k}_3)}{4k_t} \delta_{\beta\gamma} \left(2k_t^2 \varepsilon_* \frac{\dot{\phi}_\alpha^*}{H_*} + k_t^2 t_\alpha^* \left[\gamma_E + \ln \frac{k_t}{k_*} - 1 \right] - 2\varepsilon_* \frac{\dot{\phi}_\alpha^*}{H_*} \left[K^2 + \frac{k_1 k_2 k_3}{k_t} \right] + t_\alpha^* \frac{k_1 k_2 k_3}{k_t} - t_\alpha^* \left[\gamma_E + \ln \frac{k_t}{k_*} \right] \left[K^2 + \frac{k_1 k_2 k_3}{k_t} \right] \right) \quad (\text{A.50})$$

NLO from internal wavefunctions

$$\begin{aligned} \mathcal{B}_{\alpha\beta\gamma} \supseteq & \frac{(\mathbf{k}_2 \cdot \mathbf{k}_3)}{4} \frac{\dot{\phi}_\lambda^*}{H_*} \left(-\delta_{\beta\gamma} M_{\alpha\lambda}^* \left[k_t \left(1 - \ln \frac{k_t}{2k_1} \right) - 2k_1 \left(\gamma_E - 1 + \ln \frac{2k_1}{k_*} \right) + \frac{1}{k_t} (k_1 k_2 + k_1 k_3 - k_2 k_3) J_0(k_1) + \frac{k_1 k_2 k_3}{k_t^2} J_1(k_1) \right] \right. \\ & - \delta_{\alpha\lambda} M_{\beta\gamma}^* \left[k_t \left(2 - \ln \frac{k_t^2}{4k_2 k_3} \right) - 2k_2 \left(\gamma_E - 1 + \ln \frac{2k_2}{k_*} \right) - 2k_3 \left(\gamma_E - 1 + \ln \frac{2k_3}{k_*} \right) + \frac{1}{k_t} \{ k_1 k_2 + k_2 k_3 - k_1 k_3 \} J_0(k_2) \right. \\ & + \frac{1}{k_t} \{ k_2 k_3 + k_1 k_3 - k_1 k_2 \} J_0(k_3) + \left. \left. \frac{k_1 k_2 k_3}{k_t^2} \{ J_1(k_2) + J_1(k_3) \} \right] \right. \\ & + D_{\alpha\beta\gamma\lambda}^* + \frac{1}{k_t} \left(E_{\alpha\beta\gamma\lambda}^* - \frac{F_{\alpha\beta\gamma\lambda}^*}{k_t} \right) - (D_{\alpha\beta\gamma\lambda}^* + G_{\alpha\beta\gamma\lambda}^*) \left[\gamma_E + \ln \frac{k_t}{k_*} \right] \\ & \left. + L_{\alpha\beta\gamma\lambda}^* \left[k_t \left(1 - \gamma_E - \ln \frac{k_t}{k_*} \right) - \frac{k_1 k_2 k_3}{k_t^2} + \frac{1}{k_t} \left(\gamma_E + \ln \frac{k_t}{k_*} \right) \left(K^2 + \frac{k_1 k_2 k_3}{k_t} \right) \right] \right) \quad (\text{A.51}) \end{aligned}$$

where we have defined

$$\left\{ \begin{aligned} D_{\alpha\beta\gamma\lambda}^* &= 3k_t \varepsilon_* \delta_{\alpha\lambda} \delta_{\beta\gamma} + 2k_t \delta_{\beta\gamma} u_{\alpha\lambda}^* + 4k_t \delta_{\alpha\lambda} u_{\beta\gamma}^* \\ E_{\alpha\beta\gamma\lambda}^* &= \delta_{\beta\gamma} \{ k_1 (k_2 + k_3) \varepsilon_* \delta_{\alpha\lambda} - k_2 k_3 (\varepsilon_* \delta_{\alpha\lambda} + 2u_{\alpha\lambda}^*) \} \\ &\quad + \delta_{\alpha\lambda} \{ (K^2 + k_2 k_3) \varepsilon_* \delta_{\beta\gamma} - k_1 (k_2 + k_3) (\varepsilon_* \delta_{\beta\gamma} + 2u_{\beta\gamma}^*) \} \\ F_{\alpha\beta\gamma\lambda}^* &= -3k_1 k_2 k_3 \varepsilon_* \delta_{\alpha\lambda} \delta_{\beta\gamma} \\ G_{\alpha\beta\gamma\lambda}^* &= k_1 \varepsilon_* \delta_{\alpha\lambda} \delta_{\beta\gamma} - 2(k_2 + k_3) \delta_{\beta\gamma} u_{\alpha\lambda}^* - (k_1 + k_t) \delta_{\alpha\lambda} (\varepsilon_* \delta_{\beta\gamma} + 2u_{\beta\gamma}^*) \\ L_{\alpha\beta\gamma\lambda}^* &= 3\varepsilon_* \delta_{\alpha\lambda} \delta_{\beta\gamma} \end{aligned} \right. \quad (\text{A.52})$$

Relation to factorization theorem. It may appear that only Eqs. (??) and (??) satisfy the factorization theorem proved in §??. In Eqs. (??)–(??) and (??) there are divergences which do not appear to multiply local combinations of the external momenta. [As explained in §??., there is no requirement for time-dependence associated with *external* wavefunctions, as in Eqs. (??), (??) and (??), to appear in local combinations.] This is caused by the appearance in (??) of a term linear in $\partial^2 \vartheta_1 / a^2$. As noted in footnote ?? on p. ??, this term

appears because the constraints have been used to simplify the third order action. Therefore the apparent nonlocality in this term is cancelled by a delicate combination of other operators. Eqs. (??) and (??) are the lowest-order and next-order parts of this operator and combine to give a local result when symmetrized. The same is true for (??), in which $\partial^2 \vartheta_1/a^2$ appears as a proxy for the operator $\alpha_1^2 \dot{\phi}_\alpha \delta \dot{\phi}_\alpha$ which can only produce local divergences. We could have avoided this complex series of cancellations by working with an unsimplified action, at the expense of greater algebraic complexity overall.

Note that the vertex integral arising from the $\delta\phi_\alpha(\partial_i\delta\phi_\beta)^2$ operator yields no time-dependent logarithms. This is because the interaction is suppressed by spatial gradients, in agreement with the analysis of §??.